



**HAL**  
open science

**'Détection temps réel assistée par ordinateur pour le  
dépistage précoce du cancer colorectal en  
vidéocoloscopie'**

Quentin Angermann

► **To cite this version:**

Quentin Angermann. 'Détection temps réel assistée par ordinateur pour le dépistage précoce du cancer colorectal en vidéocoloscopie'. Imagerie médicale. Université de Cergy Pontoise, 2018. Français. NNT: 2018CERG0926 . tel-01928413

**HAL Id: tel-01928413**

**<https://theses.hal.science/tel-01928413v1>**

Submitted on 20 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Thèse de doctorat

---

DÉTECTION TEMPS RÉEL ASSISTÉE PAR  
ORDINATEUR POUR LE DÉPISTAGE PRÉCOCE DU  
CANCER COLORECTAL EN VIDÉOCOLOSOPHIE

---

par

QUENTIN ANGERMANN

Thèse présentée pour l'obtention du titre de docteur de l'Université Paris Seine

ETIS UMR 8051, Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS  
6 avenue du Ponceau, 95014 Cergy-Pontoise Cedex, France

Soutenue le lundi 16 juillet 2018

Devant le jury composé de :

|                    |   |                       |
|--------------------|---|-----------------------|
| Caroline PETITJEAN | MCF-HDR, LITIS, Université de Rouen           | Rapporteuse           |
| Nicolas THOME      | PU, CEDRIC, CNAM de Paris                     | Rapporteur            |
| Diane LINGRAND     | MCF, I3S, Université de Nice Sophia Antipolis | Examinatrice          |
| Michel PAINDAVOINE | PU, LEAD, Université de Bourgogne             | Examineur             |
| Aymeric HISTACE    | PU, ETIS, E.N.S.E.A.                          | Co-Directeur de thèse |
| Olivier ROMAIN     | PU, ETIS, Université de Cergy-Pontoise        | Co-Directeur de thèse |





# Table des matières

|  |           |
|--|-----------|
| <b>Introduction générale</b>   | <b>13</b> |
| Le cancer colorectal dans le monde et en France . . . . .                  | 13        |
| Mais qu'est-ce que le cancer colorectal ? . . . . .                        | 13        |
| Le dépistage du cancer colorectal . . . . .                                | 17        |
| Problématique . . . . .  | 19        |
| Plan du manuscrit . . . . .  | 19        |
| <br>   |           |
| <b>1 La détection du cancer colorectal</b>                                 | <b>21</b> |
| 1.1 Les moyens de détection du cancer colorectal . . . . .                 | 21        |
| 1.1.1 Les moyens matériels de détection du cancer colorectal . . . . .     | 21        |
| Full Spectrum Endoscopy (FUSE) . . . . .                                   | 22        |
| Narrow Band Imaging (NBI) . . . . .  | 22        |
| Fujinon Intelligent Color Enhancement (FICE) . . . . .                     | 23        |
| Coloscopie en imagerie autofluorescente (AFI) . . . . .                    | 24        |
| Chromoendoscopie . . . . .   | 25        |
| Endomicroscopie . . . . .  | 26        |
| Capsule endoscopique . . . . .   | 27        |
| Coloscopie virtuelle (Colonographie CT) . . . . .                          | 28        |
| 1.1.2 Les méthodes de détection assistée par ordinateur . . . . .          | 29        |
| Les méthodes ad-hoc . . . . .  | 30        |
| Les méthodes basées sur de l'apprentissage profond . . . . .               | 33        |
| Les méthodes hybrides . . . . .  | 34        |
| 1.2 Apprentissage machine orienté Boosting . . . . .                       | 40        |
| 1.2.1 Boosting . . . . .   | 40        |
| 1.2.2 Adaboost . . . . .   | 40        |
| 1.2.3 Cascade Adaboost . . . . .   | 41        |
| 1.3 Conclusion . . . . .   | 42        |
| <br>   |           |
| <b>2 La détection dans les images</b>                                      | <b>45</b> |
| 2.1 Une méthode pour la détection des polypes basée sur la forme . . . . . | 46        |
| 2.1.1 Constat sur la forme des polypes . . . . .                           | 46        |
| 2.1.2 La détection des formes circulaires . . . . .                        | 47        |
| La transformée de Hough . . . . .  | 47        |
| La méthode des filtres Log-Gabor . . . . .                                 | 48        |
| Conclusion intermédiaire sur la méthode à adopter . . . . .                | 49        |
| Limitation de la transformée de Hough pour la détection des polypes        | 50        |

|          |   |           |
|----------|---|-----------|
| 2.1.3    | Apprentissage machine en complément . . . . .   | 51        |
|          | Une chaîne de traitement en complément de la transformée de Hough                               | 52        |
|          | Matrice de cooccurrence et paramètres d'Haralick . . . . .                                      | 52        |
|          | Performances de l'apprentissage machine . . . . .   | 54        |
| 2.1.4    | Conclusion et limitations de cette méthode . . . . .  | 56        |
| 2.2      | L'apprentissage actif pour la détection des polypes . . . . .                                   | 58        |
| 2.2.1    | Implémentation matérielle . . . . .   | 58        |
| 2.2.2    | Paramètres d'apprentissage . . . . .  | 59        |
|          | Paramètres à motif binaire local (LBP, Local Binary Pattern) . . . .                            | 60        |
|          | Paramètres pseudo-Haar . . . . .  | 61        |
| 2.2.3    | L'algorithme de base . . . . .  | 61        |
| 2.2.4    | Résultats préliminaires sur la méthode de détection . . . . .                                   | 61        |
| 2.2.5    | L'apprentissage actif . . . . .   | 63        |
|          | Principe de l'apprentissage actif . . . . .   | 63        |
|          | Effets de l'apprentissage actif sur les performances . . . . .                                  | 65        |
|          | Conclusion sur l'apprentissage actif . . . . .  | 66        |
| 2.3      | Conclusion sur la détection dans les images . . . . .   | 66        |
| <b>3</b> | <b>La détection dans les vidéos</b>   | <b>69</b> |
| 3.1      | La détection dans des vidéos . . . . .  | 70        |
| 3.1.1    | Introduction de nouvelles métriques pour la mesure de performances<br>dans les vidéos . . . . . | 70        |
|          | Base de données vidéo . . . . .   | 72        |
|          | Protocole de mesure . . . . .   | 73        |
| 3.1.2    | Résultats préliminaires de détection dans les vidéos . . . . .                                  | 73        |
|          | Résultats . . . . .   | 74        |
| 3.2      | Amélioration des performances dans les vidéos . . . . .   | 75        |
| 3.2.1    | Fusion spatiales des zones détectées . . . . .  | 75        |
| 3.2.2    | Apprentissage actif . . . . .   | 76        |
| 3.2.3    | Cohérence temporelle . . . . .  | 78        |
|          | Fonctionnement . . . . .  | 78        |
|          | Effet de la cohérence temporelle sur la détection . . . . .                                     | 79        |
|          | Performances sans apprentissage actif . . . . .   | 79        |
|          | Temps de réaction . . . . .   | 81        |
|          | Conclusion sur la cohérence temporelle . . . . .  | 84        |
| 3.2.4    | Combinaison de l'apprentissage actif et de la cohérence temporelle .                            | 85        |
|          | Variations . . . . .  | 86        |
| 3.2.5    | Combinaison des différents types de descripteurs d'apprentissage . .                            | 87        |
| 3.2.6    | Prétraitement de l'image . . . . .  | 87        |
|          | Influence du retrait d'artefacts . . . . .  | 88        |
|          | Influence du retrait de la lumière spéculaire . . . . .   | 89        |
|          | Influence de la combinaison des deux phases de prétraitement . . . .                            | 90        |
| 3.3      | Conclusion sur la détection dans les vidéos . . . . .   | 91        |
| 3.4      | Évaluation de la robustesse . . . . .   | 93        |
| 3.5      | Détection de la présence des polypes dans les vidéos . . . . .                                  | 94        |

|   |            |
|---|------------|
| <b>4 Conclusion et Discussions</b>  | <b>97</b>  |
| 4.1 Conclusion . . . . .  | 97         |
| 4.2 Discussions . . . . .   | 99         |
| 4.2.1 Données . . . . .   | 99         |
| 4.2.2 Portage de la méthode . . . . .   | 100        |
| Cluster de Raspberry Pi3 . . . . .  | 100        |
| De l'interprété vers le compilé . . . . .   | 101        |
| Du Python vers une implémentation matérielle sur FPGA . . . . .                                   | 102        |
| 4.2.3 Tracking des polypes pour réduire le temps de calcul et améliorer<br>la détection . . . . . | 102        |
| 4.2.4 Essais cliniques . . . . .  | 103        |
| 4.2.5 La fluorescence pour caractériser les polypes . . . . .                                     | 103        |
| Définition . . . . .  | 103        |
| La fluorescence exogène . . . . .   | 104        |
| La fluorescence endogène . . . . .  | 105        |
| Liste des abréviations . . . . .  | 107        |
| Application clinique . . . . .  | 107        |
| <b>Liste des publications</b>   | <b>109</b> |
| <b>A Profils des vidéos et résultats des détections</b>   | <b>121</b> |
| <b>B Implémentation matérielle, étude comparative</b>   | <b>141</b> |
| B.1 De l'apprentissage vers l'implémentation matérielle . . . . .                                 | 141        |
| B.2 Embarquabilité . . . . .  | 141        |
| B.2.1 Temps de détection moyen sur différentes plateformes . . . . .                              | 144        |
| B.2.2 Coût des plateformes . . . . .  | 145        |
| B.2.3 Cas des processeurs ARM : accélération . . . . .  | 145        |
| B.3 Conclusion . . . . .  | 146        |
| <b>C Détection des naevus</b>   | <b>147</b> |
| C.1 Introduction . . . . .  | 147        |
| C.2 Détection et localisation des naevus . . . . .  | 149        |
| C.3 Caractérisation des naevus . . . . .  | 149        |
| C.3.1 Asymétrie . . . . .   | 149        |
| C.3.2 Bordures . . . . .  | 150        |
| C.3.3 Couleur . . . . .   | 151        |
| C.3.4 Diamètre . . . . .  | 152        |
| C.4 Performances de cette technique . . . . .   | 152        |
| C.5 Conclusion . . . . .  | 153        |



# Table des figures

|      |   |    |
|------|---|----|
| 1    | Incidence du cancer colorectal dans le monde . . . . .                        | 14 |
| 2    | Illustration du côlon humain . . . . .  | 15 |
| 3    | Classification morphologique des polypes . . . . .                            | 15 |
| 4    | Illustration concernant la forme des polypes . . . . .                        | 16 |
| 5    | Illustration de l'évolution d'un polype vers une tumeur . . . . .             | 16 |
| 6    | Illustration de la coloscopie . . . . .                                       | 17 |
| 7    | Principe de l'ESD . . . . .   | 18 |
| 1.1  | Endoscope FUSE . . . . .  | 22 |
| 1.2  | Visualisation des images avec le FUSE . . . . .                               | 22 |
| 1.3  | Fonctionnement du NBI . . . . .   | 23 |
| 1.4  | Comparaison entre endoscopie à lumière blanche et endoscopie NBI . . . . .    | 23 |
| 1.5  | Fonctionnement du FICE . . . . .  | 24 |
| 1.6  | Diagramme de Jablonski . . . . .  | 25 |
| 1.7  | Images issues de la technique AFI . . . . .                                   | 25 |
| 1.8  | Classification Kudo pour la reconnaissance de motifs . . . . .                | 26 |
| 1.9  | Illustration de l'endomicroscopie . . . . .                                   | 27 |
| 1.10 | Capsule endoscopique de la marque Pillcam . . . . .                           | 28 |
| 1.11 | Images issues d'une capsule endoscopique . . . . .                            | 28 |
| 1.12 | Comparaison entre coloscopie virtuelle et la réalité . . . . .                | 29 |
| 1.13 | Illustration de la méthode de Bernal et al. [BSFE <sup>+</sup> 15] . . . . .  | 30 |
| 1.14 | Résultat de la méthode de Bernal et al. [BSFE <sup>+</sup> 15] . . . . .      | 30 |
| 1.15 | Illustration de la méthode de Zhu et al. [ZFL10] . . . . .                    | 31 |
| 1.16 | Illustration de la méthode de Kang et al. [KD03] . . . . .                    | 31 |
| 1.17 | Illustration de la méthode de Hwang et al. [HOT <sup>+</sup> 07] . . . . .    | 31 |
| 1.18 | Résultat de la méthode de Hwang et al. [HOT <sup>+</sup> 07] . . . . .        | 32 |
| 1.19 | Illustration de la méthode de Tajbakhsh et al. [TGL16] . . . . .              | 32 |
| 1.20 | Résultat de la méthode de Tajbakhsh et al. [TGL16] . . . . .                  | 32 |
| 1.21 | Illustration de la méthode de Sevo et al. [ŠAB <sup>+</sup> 16] . . . . .     | 33 |
| 1.22 | Illustration de la méthode de Park et al. [PS16] . . . . .                    | 33 |
| 1.23 | Résultat de la méthode de Park et al. [PS16] . . . . .                        | 33 |
| 1.24 | Illustration de la méthode de Ribeiro et al. [RUH16] . . . . .                | 34 |
| 1.25 | Illustration de la méthode d'Iwahori et al. [ISH <sup>+</sup> 13] . . . . .   | 34 |
| 1.26 | Résultat de la méthode d'Iwahori et al. [ISH <sup>+</sup> 13] . . . . .       | 35 |
| 1.27 | Illustration de la méthode de Silva et al. [SHR <sup>+</sup> 14] . . . . .    | 35 |
| 1.28 | Résultat de la méthode de Silva et al. [SHR <sup>+</sup> 14] . . . . .        | 35 |
| 1.29 | Illustration de la méthode de Karkanis et al. [KIM <sup>+</sup> 03] . . . . . | 36 |

|      |  |    |
|------|--|----|
| 1.30 | Illustration de la méthode de Gross et al. [GSB <sup>+</sup> 09]                     | 36 |
| 1.31 | Illustration de la méthode d'Ameling et al. [AWP <sup>+</sup> 09]                    | 37 |
| 1.32 | Illustration de la méthode de Bae et al. [BY15]                                      | 37 |
| 1.33 | Résultat de la méthode de Bae et al. [BY15]  | 37 |
| 1.34 | Illustration du principe de détection Cascade Adaboost                               | 42 |
| 1.35 | Illustration du résultat de la détection Cascade Adaboost                            | 43 |
|      |  |    |
| 2.1  | Illustration concernant la forme des polypes   | 46 |
| 2.2  | Chaine de traitement pour la détection des polypes basée sur la forme                | 46 |
| 2.3  | Illustration de la transformée de Hough  | 47 |
| 2.4  | Illustration de la transformée de Hough circulaire                                   | 48 |
| 2.5  | Résultat de la détection basée sur les filtres Log-Gabor                             | 49 |
| 2.6  | Résultat de la détection basée sur les filtres Log-Gabor                             | 49 |
| 2.7  | Chaine de traitement pour la détection des polypes basée sur la transformée de Hough | 50 |
| 2.8  | Exemples de détections avec la transformée de Hough circulaire                       | 51 |
| 2.9  | Schéma de détection des polypes avec de l'apprentissage machine                      | 52 |
| 2.10 | Illustration du calcul de la matrice de cooccurrence                                 | 53 |
| 2.11 | Exemple d'une matrice de cooccurrence associée à son image                           | 53 |
| 2.12 | Exemples de paramètres d'Haralick  | 54 |
| 2.13 | Exemples d'images disponibles dans la base CVC-Colon                                 | 55 |
| 2.14 | Exemple de création des patches  | 55 |
| 2.15 | Schéma de détection des polypes avec de l'apprentissage Cascade Adaboost             | 56 |
| 2.16 | Exemples de détections de polypes dans des images                                    | 57 |
| 2.17 | Imagettes pour la préparation de l'apprentissage                                     | 59 |
| 2.18 | Exemple d'utilisation des LBP  | 60 |
| 2.19 | Exemples de masques proposés par Viola et Jones                                      | 61 |
| 2.20 | Exemples de masques proposés par Lienhart et Maydt                                   | 61 |
| 2.21 | Schéma de principe de l'apprentissage actif  | 64 |
| 2.22 | Évolution du $F_1$ Score avec les répétitions d'apprentissage actif                  | 65 |
| 2.23 | Résultats de la méthode avec apprentissage actif                                     | 66 |
|      |  |    |
| 3.1  | Illustration du temps de réaction  | 71 |
| 3.2  | Processus de classification d'une région d'intérêt (ROI)                             | 71 |
| 3.3  | Profil de la vidéo 18 (381 images)   | 72 |
| 3.4  | Illustration de la cohérence temporelle  | 79 |
| 3.5  | Effet de la cohérence temporelle pour la vidéo 5 (490 images)                        | 80 |
| 3.6  | Images illustrant la cohérence temporelle  | 81 |
| 3.7  | Extraits de la vidéo 7   | 83 |
| 3.8  | Présence du polype dans la vidéo 7 (440 images)                                      | 83 |
| 3.9  | Extraits de la vidéo 8   | 84 |
| 3.10 | Présence du polype dans la vidéo 8 (448 images)                                      | 84 |
| 3.11 | Score $F_1$ pour différentes combinaisons de cohérence temporelle                    | 86 |
| 3.12 | Images avec et sans suppression des bordures   | 88 |
| 3.13 | Images avec et sans lumière spéculaire   | 89 |
| 3.14 | Image sans bordure et sans lumière spéculaire  | 91 |

|      |  |     |
|------|--|-----|
| 4.1  | Configuration du cluster pour la détection basée sur Raspberry Pi3 (R.Pi3) | 101 |
| 4.2  | Principe du tracking   | 102 |
| 4.3  | Introduction à la fluorescence : diagramme de Jablonski                    | 104 |
| A.1  | Résultats de la détection pour la vidéo 1 (498 images)                     | 122 |
| A.2  | Résultats de la détection pour la vidéo 2 (773 images)                     | 123 |
| A.3  | Résultats de la détection pour la vidéo 3 (972 images)                     | 124 |
| A.4  | Résultats de la détection pour la vidéo 4 (390 images)                     | 125 |
| A.5  | Résultats de la détection pour la vidéo 5 (490 images)                     | 126 |
| A.6  | Résultats de la détection pour la vidéo 6 (857 images)                     | 127 |
| A.7  | Résultats de la détection pour la vidéo 7 (441 images)                     | 128 |
| A.8  | Résultats de la détection pour la vidéo 8 (449 images)                     | 129 |
| A.9  | Résultats de la détection pour la vidéo 9 (551 images)                     | 130 |
| A.10 | Résultats de la détection pour la vidéo 10 (840 images)                    | 131 |
| A.11 | Résultats de la détection pour la vidéo 11 (500 images)                    | 132 |
| A.12 | Résultats de la détection pour la vidéo 12 (385 images)                    | 133 |
| A.13 | Résultats de la détection pour la vidéo 13 (624 images)                    | 134 |
| A.14 | Résultats de la détection pour la vidéo 14 (2060 images)                   | 135 |
| A.15 | Résultats de la détection pour la vidéo 15 (575 images)                    | 136 |
| A.16 | Résultats de la détection pour la vidéo 16 (371 images)                    | 137 |
| A.17 | Résultats de la détection pour la vidéo 17 (797 images)                    | 138 |
| A.18 | Résultats de la détection pour la vidéo 18 (381 images)                    | 139 |
| B.1  | Temps de détection moyens par image pour chacune des plateformes           | 144 |
| B.2  | Rapport performances - prix des différentes plateformes                    | 145 |
| B.3  | Temps de détection avec et sans NEON sur Altera DE1 SoC                    | 146 |
| C.1  | Chaîne de traitement des naevus  | 148 |
| C.2  | Exemple de détections et comparaison à la vérité terrain                   | 149 |
| C.3  | Asymétrie : Prétraitement pour déterminer la symétrie                      | 150 |
| C.4  | Asymétrie : Recherche de similarités entre deux demi naevus                | 150 |
| C.5  | Bordures : Détermination de contours                                       | 151 |
| C.6  | Bordures : Comparaison des longueurs des contours                          | 151 |
| C.7  | Couleur : Illustration d'un naevus recadré                                 | 152 |
| C.8  | Couleur : Matrice de cooccurrence associée à la figure C.7                 | 152 |
| C.9  | Diamètre : Formule pour déterminer le diamètre réel de l'objet             | 152 |



# Liste des tableaux

|      |  |    |
|------|--|----|
| 1.1  | Tableau comparatif du panel exhaustif des méthodes existantes pour la détection des polypes colorectaux . . . . .  | 39 |
| 1.2  | Algorithme Adaboost . . . . .  | 41 |
| 1.3  | Algorithme Cascade Adaboost . . . . .  | 43 |
| 2.1  | Performances de la détection des polypes . . . . .   | 50 |
| 2.2  | Performances de la classification sur des patches de polypes et non des images de polypes . . . . .  | 56 |
| 2.3  | Algorithme pour la détection des polypes dans les images . . . . .   | 62 |
| 2.4  | Résultats de la détection sur images fixes . . . . .   | 63 |
| 2.5  | Principe de l'apprentissage actif . . . . .  | 64 |
| 2.6  | Résultats de la détection sur images fixes avec l'apprentissage actif (A.A.) .   | 65 |
| 3.1  | Niveaux de difficulté de détection des polypes dans les vidéos . . . . .   | 73 |
| 3.2  | Algorithme de détection des polypes dans les vidéos . . . . .  | 74 |
| 3.3  | Résultats préliminaires de la détection dans les vidéos sans optimisations .   | 74 |
| 3.4  | Algorithme de détection des polypes dans les vidéos avec fusion spatiale . .   | 76 |
| 3.5  | Comparaison des performances de la méthode avec fusion spatiale pour la classification utilisant les motifs locaux binaires (LBP) . . . . .              | 76 |
| 3.6  | Comparaison des performances de la méthode avec fusion spatiale pour la classification utilisant les paramètres pseudo-Haar . . . . .                    | 76 |
| 3.7  | Résultats de la détection dans les vidéos avec l'apprentissage actif pour la classification utilisant les motifs binaires locaux (LBP) . . . . .         | 77 |
| 3.8  | Résultats de la détection dans les vidéos avec l'apprentissage actif pour la classification utilisant les paramètres pseudo-Haar . . . . .               | 77 |
| 3.9  | Algorithme de détection des polypes dans les vidéos . . . . .  | 78 |
| 3.10 | Influence de la cohérence temporelle sur les résultats utilisant la classification basée sur les LBP sans apprentissage actif . . . . .                  | 80 |
| 3.11 | Influence de la cohérence temporelle sur les résultats utilisant la classification basées les paramètres pseudo-Haar sans apprentissage actif . . . . .  | 80 |
| 3.12 | Influence de la cohérence temporelle sur le temps de réaction avec la classification utilisant les paramètres pseudo-Haar sans apprentissage actif . . . | 82 |
| 3.13 | Temps de réaction pour chaque vidéo sans cohérence temporelle . . . . .  | 82 |
| 3.14 | Temps de réaction pour chaque vidéo avec cohérence temporelle . . . . .  | 82 |
| 3.15 | Résultats de la détection dans les vidéos avec la cohérence temporelle et l'apprentissage actif . . . . .  | 85 |
| 3.16 | Comparaison entre classifications classiques et combinaison de classifications   | 87 |

|      |   |     |
|------|---|-----|
| 3.17 | Comparaisons des performances de détection entre images normales et images détournées . . . . .                                 | 88  |
| 3.18 | Comparaisons des temps de traitement entre images normales et images détournées . . . . .                                       | 89  |
| 3.19 | Comparaisons des performances de détection entre images normales et images sans lumière spéculaire . . . . .                    | 90  |
| 3.20 | Comparaisons des temps de traitement entre images normales et images détournées . . . . .                                       | 90  |
| 3.21 | Comparaisons des performances de détection entre images normales et images détournées . . . . .                                 | 91  |
| 3.22 | Résumé des gains en performances suivant les optimisations . . . . .  | 93  |
| 3.23 | Résultats de la détection sur une nouvelle base de données vidéos . . . . .   | 94  |
| 3.24 | Comparaison des résultats de la détection sur les deux bases de données vidéos . . . . .  | 95  |
| 3.25 | Comparaison des résultats de la détection de la présence des polypes dans les vidéos de la base [BHM <sup>+</sup> 18] . . . . . | 96  |
| 4.1  | Liste des fluorophores exogènes . . . . .   | 105 |
| 4.2  | Liste des fluorophores endogènes . . . . .  | 106 |
| B.1  | Compatibilité avec le temps réel des systèmes dans la littérature . . . . .   | 142 |
| C.1  | Méthode de détection des naevus . . . . .   | 148 |
| C.2  | Performances de la méthode de détection et caractérisation des naevus . . .   | 153 |

# Introduction générale

## Le cancer colorectal dans le monde et en France

Le cancer colorectal est une des principales causes de décès par cancer dans le monde (figure 1), avec une incidence estimée à 1 370 600 nouveau cas en 2012. Il est fatal dans la moitié des cas (693 000 décès estimés en 2012) [Soc15]. Dans les pays développés, c'est un des cancers les plus fréquents et meurtriers (les autres les plus fréquents sont le cancer du sein et le cancer des poumons), avec 736 900 nouveaux cas déclarés en 2012.

En France, le cancer colorectal est le cancer le plus fréquent derrière le cancer du sein et de la prostate (40 000 nouveaux cas par an). L'incidence de ce cancer n'a fait qu'augmenter durant ces 20 dernières années. Il est la deuxième cause de mortalité par cancer juste après celui du poumon. En 2015, selon l'Institut National du Cancer, le cancer colorectal a provoqué 17 500 décès, soit 1 décès toutes les demi-heures. Cependant, si le cancer est détecté suffisamment tôt, il peut se guérir dans 9 cas sur 10. D'où l'importance du dépistage précoce car c'est le moyen le plus efficace pour lutter contre cette maladie.

## Mais qu'est-ce que le cancer colorectal ?

Anatomiquement parlant, le cancer colorectal, comme son nom l'indique, apparaît dans le côlon et le rectum (qui forment en partie le gros intestin). Il est intéressant de noter que ces deux éléments peuvent encore se diviser en 5 éléments, en particulier pour le côlon qui se sépare en 4 sections distinctes (figure 2) :

- Le côlon ascendant (ou droit)
- Le côlon transverse
- Le côlon descendant (ou gauche)
- Le côlon sigmoïde

En effet, selon sa localisation, le cancer colorectal peut prendre différentes formes. Cette maladie se développe lentement, sur une période d'environ 10 à 20 ans. Dans la plupart des cas, ce sont des tumeurs bénignes qui apparaissent dans le côlon, qu'on appelle polypes (figure 4), très fréquentes, et qui pour certaines de ces lésions, représentent le stade précoce du cancer colorectal.

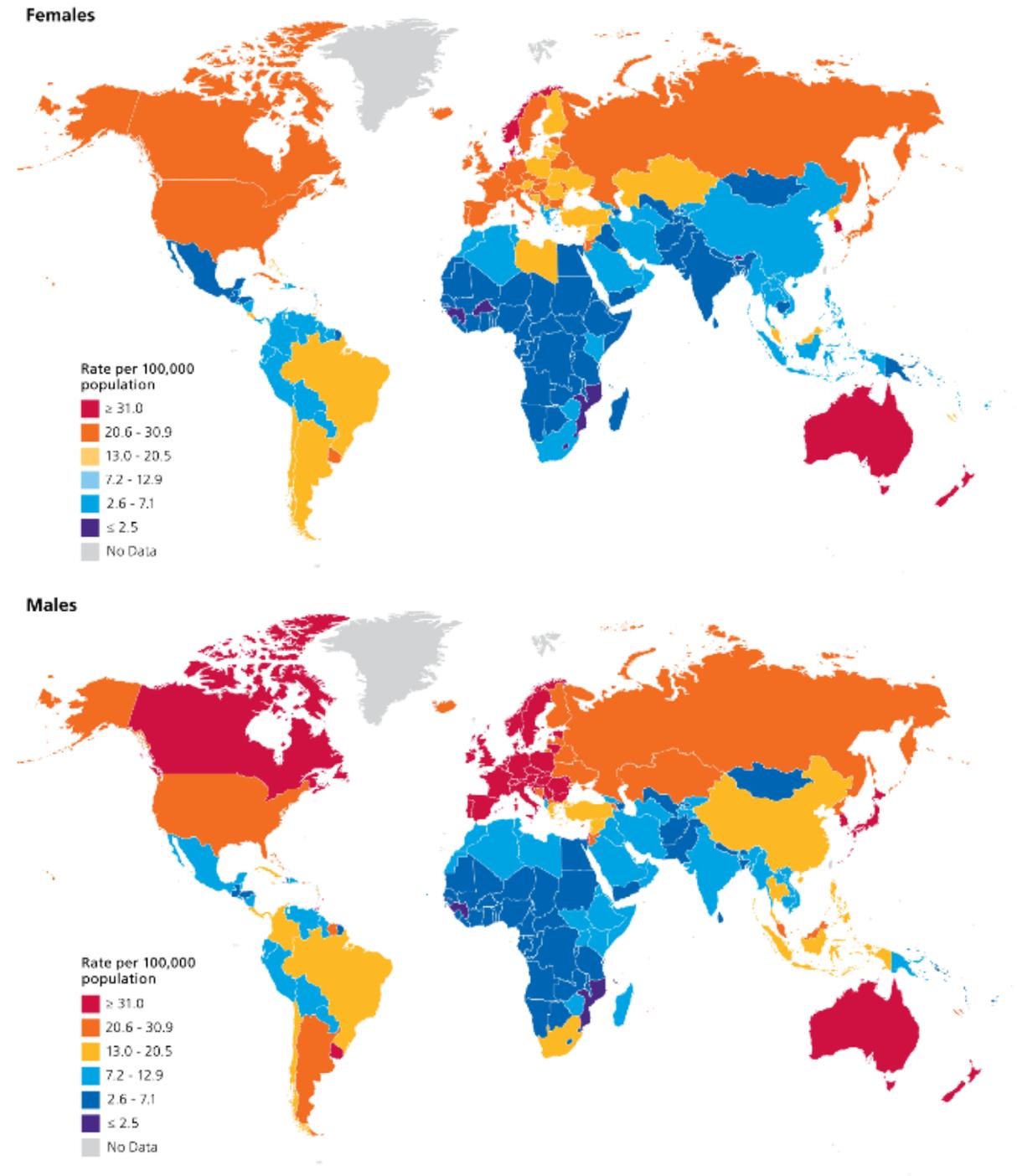


FIGURE 1 – Incidence du cancer colorectal dans le monde [Soc15]

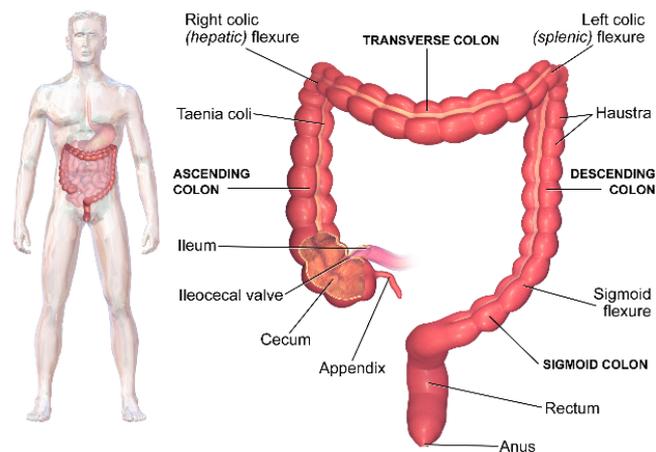


FIGURE 2 – Illustration du côlon humain (Source : Wikimedia Commons)

Ces polypes peuvent être de différents types :

- pédiculés (lésion rattachée à la paroi du côlon par un pied, comme un champignon) (figure 3.a)
- sessiles (lésion prenant la forme du sommet d'une colline posé sur la paroi) (figure 3.b)
- plans (lésion presque imperceptible en endoscopie standard) (figures 3.c, 3.d, 3.e, 3.f)

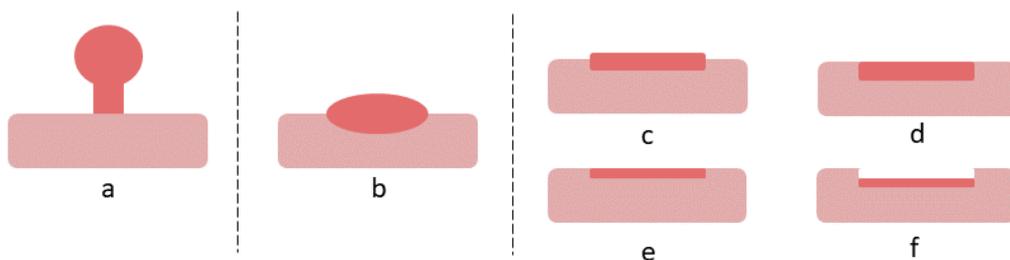


FIGURE 3 – Classification morphologique des polypes

Ces polypes peuvent être de nature bénigne ou maligne. Lorsqu'on détecte chez un patient plus d'une dizaine de polypes, on parle alors de polypose. Du point de vue histologique (qui permet d'explorer la structure du polype et d'en trouver sa dangerosité), il existe quatre variétés de polypes :

- Les polypes adénomateux : il s'agit de polypes qui peuvent se développer en cancer (la figure 5 montre l'évolution d'un polype adénomateux en tumeur).

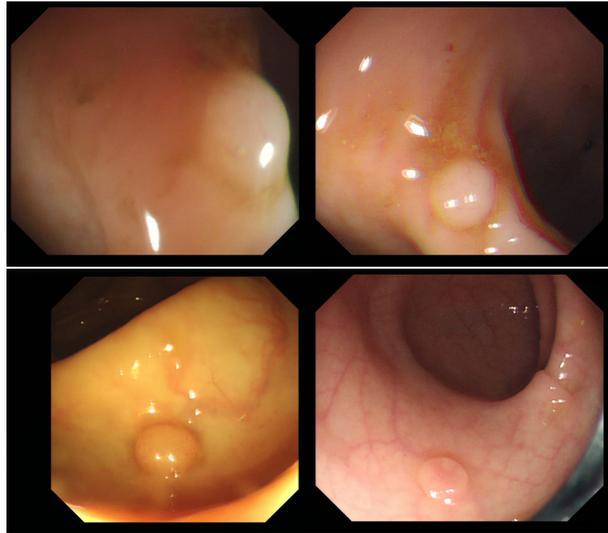


FIGURE 4 – Illustration concernant la forme des polypes

- Les polypes hyperplasiques : il s’agit de polypes qui apparaissent naturellement et qui ne dégénèrent pas nécessairement en adénomes. Cependant, un nombre élevé de ces lésions peut traduire une grande probabilité de développer des polypes adénomateux.
- Les polypes juvéniles : il s’agit de polypes qui sont des kystes d’origine inflammatoire.
- Les pseudo-polypes inflammatoires : il s’agit de polypes qui résultent d’une cicatrisation due à l’ulcération du côlon ou de la maladie de Crohn.

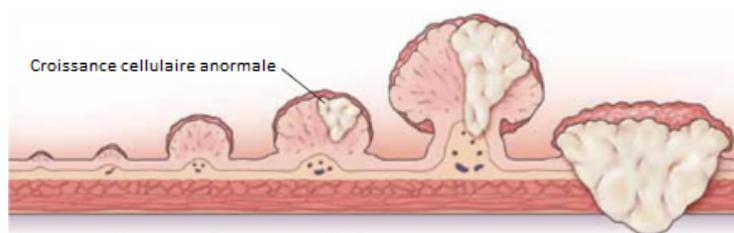


FIGURE 5 – Illustration de l’évolution d’un polype vers une tumeur (Source : Atlanta Center for Gastroenterology)

Le cancer colorectal a de plus grandes probabilités d’apparaître chez l’humain dès l’âge de 50 ans. C’est pour cela que dans la plupart des pays développés et plus particulièrement en France, les personnes de plus de 50 ans sont encouragées à passer un test de dépistage.

## Le dépistage du cancer colorectal

En France, le cancer colorectal est dépisté en plusieurs étapes. La première consiste à réaliser un test immunologique (depuis 2015) qui permet de rechercher des traces de sang dans les selles du patient. Ce test analyse la présence de sang résultant de micro-saignements dus à la possible présence de polypes. Si le test est positif (4% des cas d'après l'Assurance Maladie), un examen plus poussé est demandé : la vidéo coloscopie, qui est la référence clinique en matière de détection et d'ablation des polypes. Dans certains cas particuliers (antécédents familiaux, syndrome de Lynch, polypose adénomateuse familiale, maladies inflammatoires), on demande aux personnes de faire plus fréquemment cet examen.

La coloscopie (figure 6) est basée sur l'utilisation d'un coloscope constitué d'une source de lumière au xénon, un système optique (fibre optique ou caméra vidéo) et avec un (ou plusieurs) canal opérateur. Ce canal permet au médecin d'insuffler de l'air, de nettoyer le côlon ou d'envoyer de l'eau. Ils permettent aussi d'insérer des clips pour effectuer des prélèvements de la muqueuse (biopsies) ou d'utiliser des instruments chirurgicaux. Quand un polype est retiré, il est automatiquement envoyé en laboratoire d'histopathologie pour en faire l'étude au niveau cellulaire.

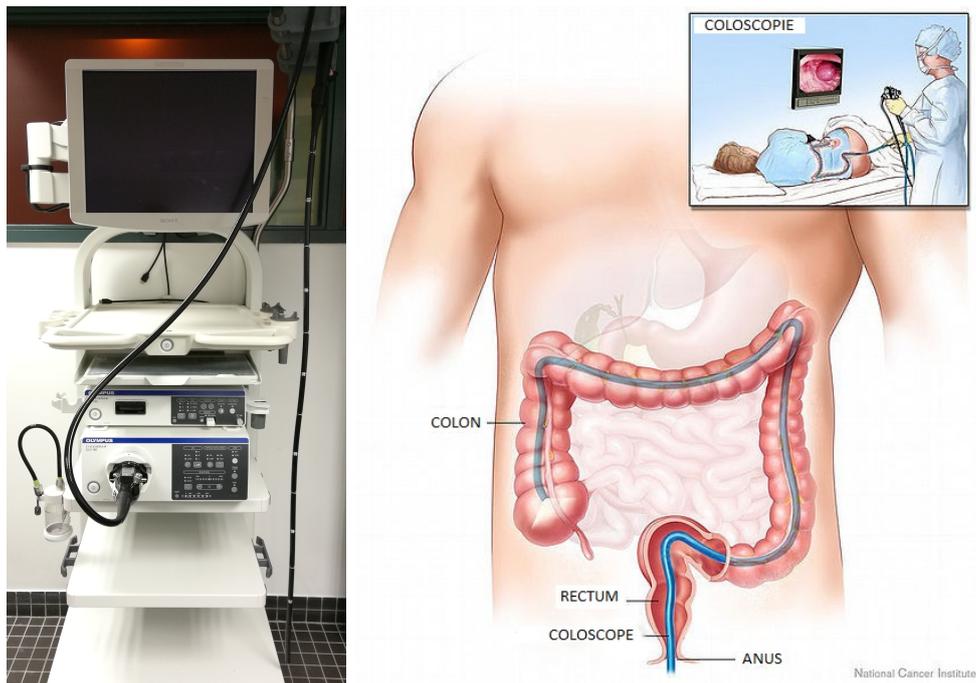


FIGURE 6 – Illustration de la coloscopie, à gauche une colonne d'endoscopie digestive, à droite une illustration de l'examen (Source : National Cancer Institute)

A l'autre extrémité du coloscope, on trouve une unité de traitement avec la source de lumière, un ordinateur et une console de visionnage. Cette unité permet au médecin de voir ce qu'il fait en temps réel et d'enregistrer éventuellement une partie des examens.

L'imageur fonctionne donc en temps réel afin de fournir l'image au médecin le plus vite possible dans le but de ne pas avoir de décalage entre ce qui est vu et ce qu'il fait à un instant  $t$ .

L'examen lui-même se déroule sous anesthésie générale après avoir préparé le côlon (grâce à une alimentation pauvre en résidus plusieurs jours avant l'examen et en buvant une solution, le polyéthylèneglycol, la veille). En pratique, on suit le protocole suivant : le coloscope est inséré par l'anus puis est doucement glissé jusqu'à l'intestin, le coloscope étant géré par le gastroentérologue avec un système de leviers. Puis il insuffle de l'air pour éloigner les parois et avancer avec prudence. Le but est d'atteindre le cecum, à la limite de la valve ileocecale (figure 2). La visualisation de la paroi du côlon sera effectuée lors du retrait du coloscope. Durant cette visualisation, le médecin cherche des polypes. A chaque fois qu'un polype est détecté, le médecin va le retirer en utilisant la méthode ESD (Endoscopic Submucosal Dissection ou dissection sous-muqueuse) décrite sur la figure 7 si la taille du polype le permet sans danger pour le patient.

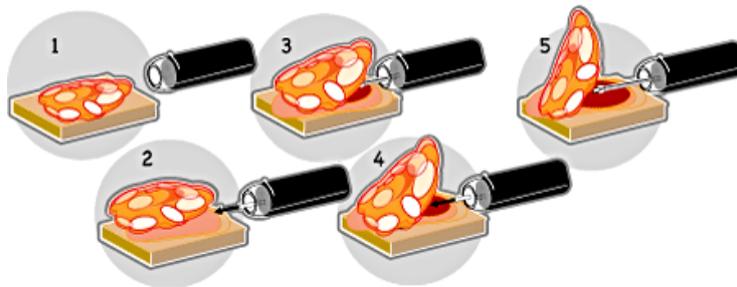


FIGURE 7 – Principe de l'ESD, 1 Marquage de la lésion, 2 Injection sous muqueuse (d'une solution saline), 3 incision sous muqueuse circonférentielle, 4 et 5 dissection progressive de la lésion (image tirée du site internet du Service de Gastroentérologie de l'hôpital Cochin, Assistance Publique - Hôpitaux de Paris)

Une fois le polype retiré, il est envoyé en anatomo-pathologie afin d'être analysé et de connaître la nature du polype, c'est-à-dire savoir s'il s'agit d'un adénome ou non.

Sur l'aspect matériel, on note une migration du matériel vers la haute définition, qui permet d'obtenir un meilleur taux de détection des adénomes (28.8% en HD contre 24.3% en SD) et des polypes (42.2% en HD contre 37.8% en SD)[DMS14]. De plus, la haute définition permet d'avoir moins d'artefacts et l'aspect de la paroi du côlon est plus naturel.

Cependant, il a été montré que la variabilité inter et intra médecins peut monter jusqu'à 26% dans la tâche de détection visuelle des polypes durant un examen [RHP+10]. Ceci s'explique par plusieurs facteurs tels que :

- l'expérience du praticien,
- l'heure de l'examen,
- la qualité de préparation du côlon,

- des polypes cachés (derrière les plis du côlon par exemple),
- la taille des polypes (en particulier les plus petits).

## Problématique

Comme on a pu le voir précédemment, le cancer colorectal est un enjeu de santé publique majeur dans les pays développés. Sa réduction de taux de mortalité passe ainsi par un meilleur dépistage, mais surtout par la réduction du taux de polypes manqués par les médecins lors de l'examen clinique de coloscopie. Afin de réduire ce taux de polypes non détectés, on se propose de mettre au point dans cette thèse un nouvel outil permettant d'aider le médecin lors des coloscopies. L'outil que l'on va mettre au point se doit de respecter les contraintes suivantes :

- être capable de détecter les polypes en temps réel (c'est-à-dire que le temps pour traiter une image issue de coloscopie doit être inférieur à 40 millisecondes)
- être compatible avec du matériel dont les ressources de calculs sont standards et limitées
- obtenir des performances élevées en termes de sensibilité et de spécificité
- être compatible avec l'utilisation en routine clinique (et donc ne pas modifier la pratique clinique du médecin) et alerter le médecin sur la présence de polypes durant l'examen.

## Plan du manuscrit

Pour répondre à cette problématique, on s'appuiera sur plusieurs chapitres permettant de mettre en lumière les différentes étapes du travail réalisé :

Nous proposons dans le chapitre 1 de présenter l'état de l'art en matière d'outils permettant d'aider le médecin à mieux détecter les polypes. Ce chapitre permet de mettre en évidence les limitations actuelles et le positionnement à adopter pour développer notre outil ainsi que la méthodologie permettant l'évaluation de ses performances.

Le chapitre 2 se consacrera à l'outil développé pour la détection dans les images issues de coloscopies, et en particulier sur les pistes explorées et adoptées afin de développer un outil capable de répondre aux différentes contraintes imposées par une utilisation en routine clinique.

Le chapitre 3 permettra de montrer l'adaptation de notre méthode à la détection dans les vidéos issues de coloscopies. Adaptation qui se rapproche de cas réels et donc de l'utilisation clinique de notre outil.

Enfin, une conclusion globale synthétisera les contributions de ce travail et apportera les éléments de discussions afin d'identifier les pistes principales qui permettront d'amé-

liorer l'approche proposée.

# Chapitre 1

## La détection du cancer colorectal en vidéocoloscopie

Comme vu précédemment, le médecin, lorsqu'il effectue l'examen, peut être amené à manquer jusqu'à 26% des polypes. C'est dans cet optique de réduire ce taux que plusieurs méthodes et outils de détection ont été mis au point.

Dans ce chapitre, on va se consacrer à présenter les techniques mises au point afin de réduire ce taux de polypes manqués qui sont aussi diverses que variées. De plus, on va aussi en déduire et présenter quels peuvent être les outils que l'on peut utiliser pour mettre au point notre méthode permettant de détecter les polypes de manière automatique, en respectant les contraintes présentées précédemment.

### 1.1 Les moyens de détection du cancer colorectal

Parmi tous les moyens qui existent afin de lutter contre ce problème, on peut distinguer deux grandes familles :

- Les outils matériels, qui consistent à proposer un nouvel outil physique (comme un nouveau coloscope), avec des propriétés spécifiques qui permettent d'améliorer la détection des polypes.
- Les outils d'aide à la détection assistée par ordinateur, qui consistent à détecter durant la coloscopie (ou dans des vidéos de coloscopies) les polypes afin d'aider le médecin dans son analyse.

On va donc s'intéresser maintenant à voir ce que proposent ces deux grandes familles, en termes d'améliorations et de limitations.

#### 1.1.1 Les moyens matériels de détection du cancer colorectal

Outre la vidéocoloscopie (SD ou HD) qui est la méthode de détection la plus répandue, d'autres moyens matériels ont été mis en place afin d'améliorer les taux de détection de

polypes et d'adénomes. Dans les paragraphes suivants, sont présentés ceux ayant eu un impact réel sur le sujet et commercialisés aujourd'hui.

### Full Spectrum Endoscopy (FUSE)

Basé sur un endoscope modifié, le FUSE se distingue par l'ajout de 2 caméras de chaque côté du coloscope (figure 1.1). Cela permet au médecin de voir le côlon à 330 degrés [DMS14] et donc de voir normalement plus facilement la paroi du côlon, et bien sûr, le canal opérateur reste disponible afin d'effectuer des biopsies ou de procéder au retrait de polypes et d'adénomes.



FIGURE 1.1 – Endoscope FUSE

La colonne d'endoscopie est, elle, équipée de 3 écrans qui retranscrivent l'image des 3 caméras (figure 1.2).

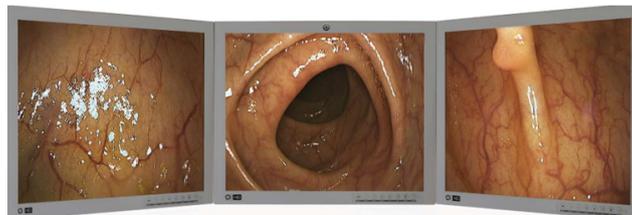


FIGURE 1.2 – Visualisation des images avec le FUSE

Comparée à la coloscopie SD à lumière blanche, il y a une meilleure détection des polypes durant l'examen (le taux d'adénomes manqués chute de 41% à 7% en utilisant le FUSE) [GSH<sup>+</sup>14]. Cependant, cette technique nécessite que le médecin apprenne à se servir de ce nouvel outil, ce qui change le protocole opératoire, et représente donc un point faible de cette méthode en terme d'utilisabilité.

### Narrow Band Imaging (NBI)

Le Narrow Band Imaging (NBI, ou imagerie à bande étroite en français) est une technique de coloscopie, développée par Olympus, qui met l'accent sur les longueurs d'ondes

bleue (415 nm) et verte (540 nm) pour améliorer certains détails de la muqueuse [PRW11]. Cette accentuation est obtenue par des filtres optiques qui n'autorisent que les lumières bleue et verte à passer dans l'endoscope. Ce filtre s'active sur simple appui d'un bouton, ce qui permet au médecin de basculer entre le mode classique (lumière blanche) au mode NBI. La figure 1.3 montre le principe du NBI depuis la source lumineuse jusqu'au capteur de l'imageur. Des études ont montré que le taux de polypes manqués chute de 57% à 31% en comparaison à l'endoscopie en lumière blanche. Il en est de même pour le taux d'adénomes manqués qui lui passe de 49% à 27% [GBC<sup>+</sup>11]. Cependant, contrairement à la coloscopie en lumière blanche, l'intensité lumineuse du NBI est plus faible (figure 1.4) et rend donc la lecture moins aisée.

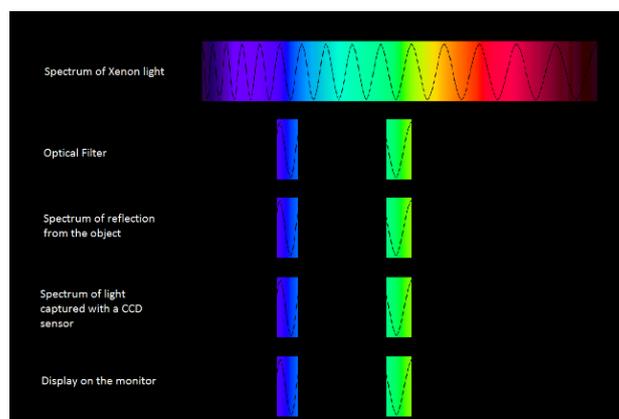


FIGURE 1.3 – Fonctionnement du NBI

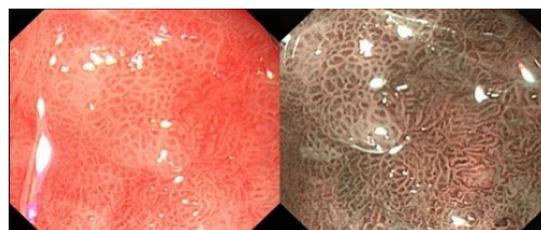


FIGURE 1.4 – Comparaison entre endoscopie à lumière blanche et endoscopie NBI

### Fujinon Intelligent Color Enhancement (FICE)

Le FICE utilise un endoscope classique avec une source de lumière au Xénon [Fuj07] qui possède un spectre lumineux s'étendant entre 400 et 700 nm. La lumière part de l'endoscope et est réfléchi sur la paroi du côlon qui est récupérée par un capteur CCD (charge-coupled device, dispositif à transfert de charge en français). Puis l'information du capteur est convertie en image. Contrairement au NBI, une étape de traitement d'images est effectuée en sélectionnant des filtres virtuels afin de faire apparaître des zones intéressantes du côlon. Le FICE permet de sélectionner les images spectrales avec les longueurs

d'ondes appropriées pour maximiser la différence dans le spectre de réflexion. Cela permet de reconstruire des images à contraste élevé pour une grande variété de tissus dans le corps. Le FICE effectue le traitement d'image (application des filtres) et le médecin peut passer facilement de la coloscopie classique au mode FICE. La figure 1.5 montre le principe du FICE depuis la source lumineuse jusqu'au capteur de l'imageur.

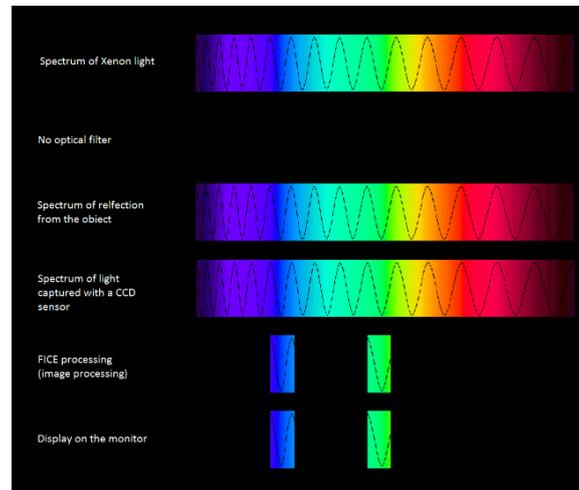


FIGURE 1.5 – Fonctionnement du FICE

### Coloscopie en imagerie autofluorescente (AFI)

Basée sur un coloscope modifié, cette technique fait appel au principe d'autofluorescence. Le principe de fluorescence est le suivant : quand la lumière réagit avec la matière, des photons sont absorbés. Il y a alors une transition électronique depuis le plus bas niveau d'énergie  $S_0$  vers des niveaux plus élevés  $S_1$ ,  $S_2$  ou  $S_n$ . Les niveaux les plus hauts sont atteints si l'énergie est assez élevée. Ensuite intervient une reconversion de l'énergie interne qui permet de revenir au niveau  $S_1$   $V_n$ . On passe alors du niveau  $V_n$  au niveau  $V_0$  grâce à des interactions entre des molécules de fluorophore et de solvant. Et ce n'est seulement qu'au moment de la transition depuis le niveau  $S_1$   $V_0$  qu'intervient le phénomène de fluorescence [Alb01] (voir figure 1.6).

Techniquement parlant, la fluorescence est l'émission de photons depuis un état singulier  $S_1$ . L'intensité du spectre de fluorescence dépend de la population de molécules excitées à l'état  $S_1$   $V_0$ .

L'électron suit un circuit fermé : il commence à  $S_0$   $V_0$  vers l'état  $S_n$   $V_m$ , puis par reconversion interne, va à l'état  $S_1$   $V_0$ . Ensuite, il passe à l'état  $S_0$   $V_0$  avec une émission de photons (la fluorescence). Mais la fluorescence n'est pas permanente, elle a une durée de vie qui varie de la nanoseconde à la picoseconde. Il existe par ailleurs deux types de fluorescence : l'endogène et l'exogène [AEW92].

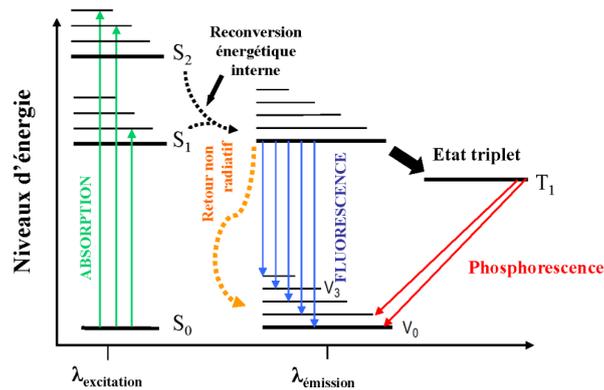


FIGURE 1.6 – Diagramme de Jablonski

L'autofluorescence est aussi appelée fluorescence endogène. Elle est basée sur des fluorophores naturels qui sont présents dans le corps du patient et certains de ces fluorophores peuvent révéler la présence de cellules cancéreuses.

Côté technique, la source de lumière blanche est séparée en lumière d'excitation (à une longueur d'onde particulière, comprise entre 400 et 500 nm) et en lumière verte. Puis, en récupérant la lumière réfléchie, une image est affichée à l'écran [ITA<sup>+</sup>13](voir figure 1.7).

L'avantage de cette technique est de permettre une meilleure détection des lésions [RHP<sup>+</sup>10] [MFS<sup>+</sup>12]. Cependant, elle peut aussi révéler des zones sans lésions et entraîne alors de nombreuses fausses détections [RKH14].

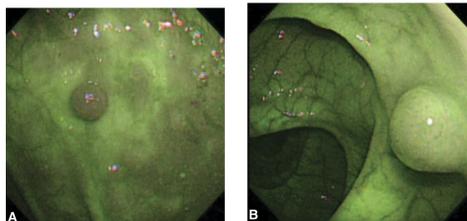


FIGURE 1.7 – Images issues de la technique AFI

## Chromoendoscopie

La chromoendoscopie vise à améliorer la précision dans le diagnostic lors de l'endoscopie gastro-intestinale, en utilisant différents colorants. En particulier, deux types de colorants sont utilisés en chromoendoscopie :

- Des colorants de surface qui ne réagissent pas avec les cellules de la muqueuse mais qui accentuent le relief pour optimiser le diagnostic des lésions (par exemple, le carmin d'indigo)

- Des colorants essentiels qui réagissent spécifiquement avec certaines cellules de la muqueuse et permettent de deviner leur nature (par exemple, le bleu de méthylène)

La chromoendoscopie utilisant le carmin d'indigo est la méthode d'exploration de la muqueuse la plus utilisée. Cette méthode permet de définir plus précisément les irrégularités de relief et surface des lésions. Le colorant est appliqué directement sur la muqueuse avec un cathéter durant l'examen [RKH14]. Puis le médecin compare le motif de la zone d'intérêt à la classification Kudo (figure 1.8)[KTN<sup>+</sup>96].

Mais avant d'effectuer une chromoendoscopie, une coloscopie classique est requise. Elle ne sera utilisée que sur des zones suspectes ou des lésions planes dans l'optique de définir les bords de la lésion détectée par coloscopie. Ainsi, les biopsies peuvent être mieux effectuées et les polypectomies ont une meilleure précision. La chromoendoscopie permet également de faire une reconnaissance des motifs afin de deviner l'histologie de la lésion.

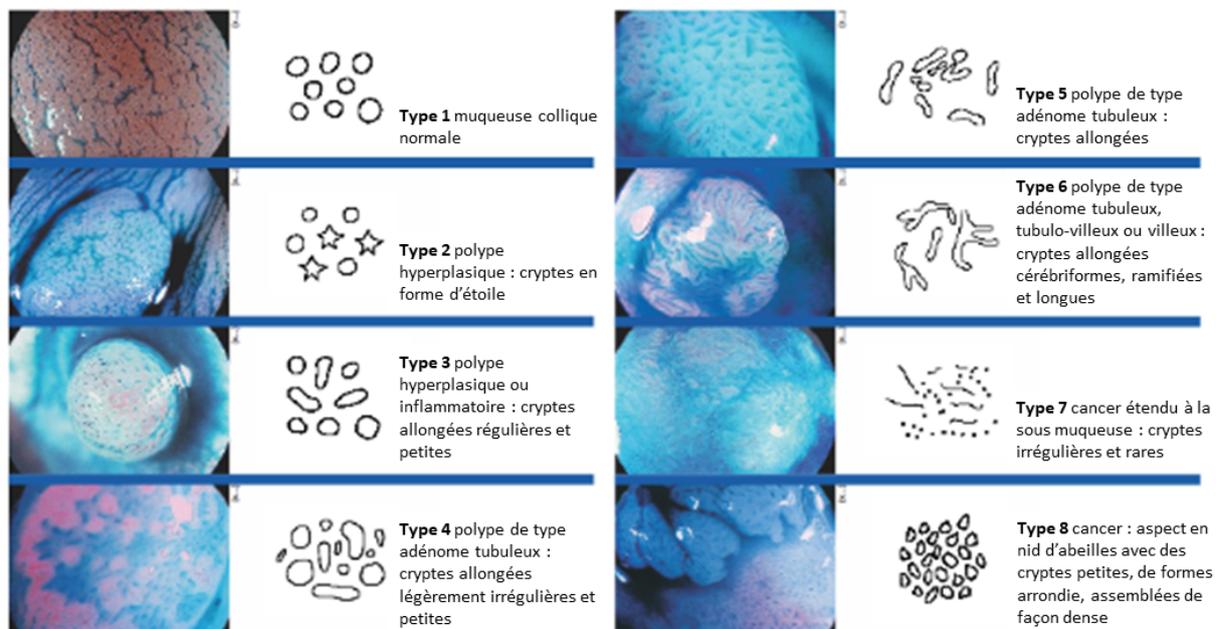


FIGURE 1.8 – Classification Kudo pour la reconnaissance de motifs

La chromoendoscopie est donc utile pour la détection des adénomes plans et pour mieux définir les contours d'une lésion [LHN<sup>+</sup>06] [BB10]. Cependant, elle ne doit pas être effectuée avant une coloscopie conventionnelle. Finalement, elle peut être utile dans le diagnostic des lésions mais une biopsie est toujours nécessaire afin d'effectuer une analyse histologique.

## Endomicroscopie

L'endomicroscopie est un examen coloscopique basé sur un endomicroscope qui est soit un microscope miniaturisé couplé à un endoscope, soit un microscope intégré à un endo-

scope spécifique. Le microscope utilise un laser à argon (488 nm) [RKH14] et fonctionne avec l'ajout d'un agent de contraste injecté au patient (fluoresceine). Il permet au médecin de faire une biopsie virtuelle durant un examen (figure 1.9). Le but de cette méthode est de faire un diagnostic rapide des lésions. Des études récentes ont montré la très bonne précision dans le diagnostic : 80% à 95% de détection [LYS<sup>+</sup>13].

Cependant, cet examen requiert la présence à la fois d'un hépatogastroentérologue et d'un histologiste afin de déterminer la dangerosité d'une lésion. Et l'agent de contraste peut provoquer une coloration jaune de la peau pour le patient [KBV<sup>+</sup>04].

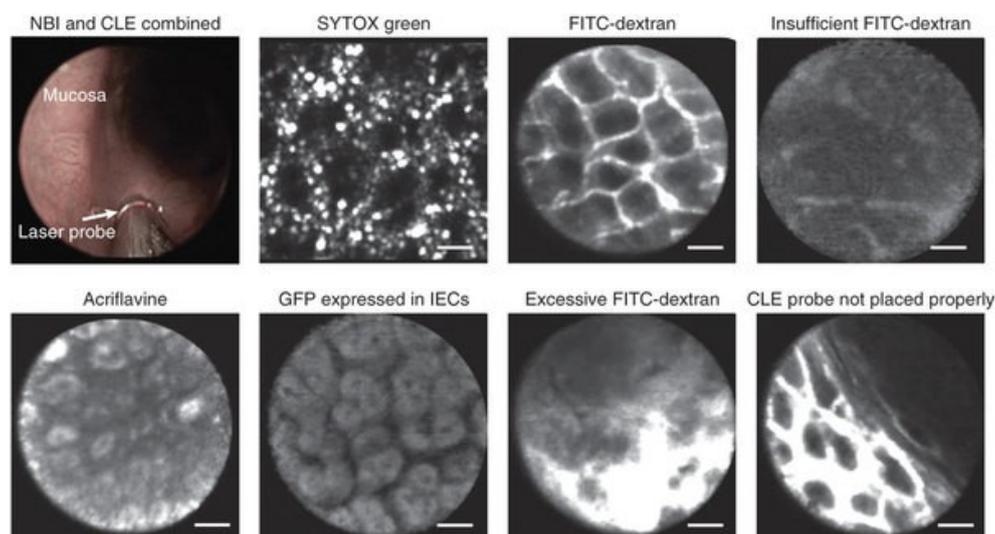


FIGURE 1.9 – Illustration de l'endomicroscopie (Source : Society for Mucosal Immunology)

### Capsule endoscopique

La capsule endoscopique est une méthode d'endoscopie utilisant une capsule avec de l'électronique embarquée (figure 1.10). Elle est constituée d'un ou deux capteurs vidéo pour enregistrer les images du côlon, d'un processeur de traitement des données et une antenne radio-fréquence pour envoyer les images à un récepteur externe [AHR<sup>+</sup>15] (datalogger). Le récepteur externe est relié à un jeu d'antennes qui sont collées sur le corps du patient pour récupérer les données de la capsule et les stocker dans le datalogger.

L'examen se déroule d'une façon totalement différente de la coloscopie. Après un nettoyage du côlon grâce à une solution à base de polyéthylèneglycol (comme en coloscopie), le patient ingère la capsule sous la supervision du médecin puis peut retourner à sa routine habituelle. Pendant ce temps, la capsule prend des photos du côlon (figure 1.11, au rythme de 4 à 35 images par seconde pendant environ 8 heures). Une fois l'examen terminé, le médecin récupère le boîtier récepteur afin de regarder toutes les images pour chercher la présence de polypes ou d'adénomes (à titre d'exemple, une capsule endoscopique prend en moyenne 150 000 images que le médecin analyse en 1h et 30 minutes). Il s'agit donc



FIGURE 1.10 – Capsule endoscopique de la marque Pillcam

ici d'un examen moins contraignant que la coloscopie [WBB<sup>+</sup>13].

Cette méthode possède une sensibilité acceptable pour la détection des polypes grâce aux capsules de seconde génération (plus de 80%) [SHMN<sup>+</sup>11]. Cependant, l'image obtenue est de moins bonne qualité qu'en vidéocoloscopie (à cause de la résolution qui est de 500 pixels par 500 pixels). Et comme la capsule n'est pas "intelligente", le temps de post-traitement des données par le médecin est long. De plus, il est impossible d'effectuer une biopsie ou de retirer les lésions suspectes une fois détectées : il faut alors effectuer une coloscopie.

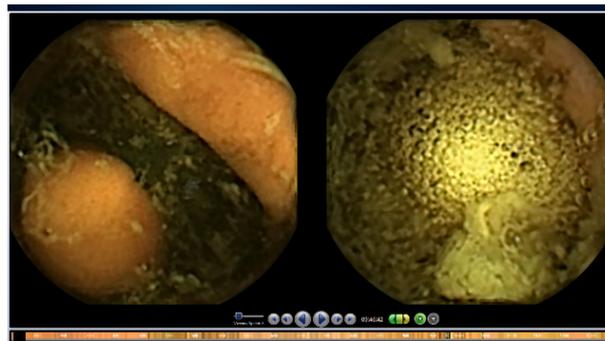


FIGURE 1.11 – Images issues d'une capsule endoscopique

### Coloscopie virtuelle (Colonographie CT)

Contrairement à toutes les méthodes précédentes, la coloscopie virtuelle ne demande pas un endoscope ou une capsule. Elle fonctionne avec un scanner ou une IRM. L'étude

du côlon est effectuée grâce à la reconstruction virtuelle en 3D du côlon à partir des données acquises par le scanner. Cette méthode évolue donc avec les améliorations dans le domaine informatique [BfCCS<sup>+</sup>10]. Afin d'effectuer cet examen, on injecte un agent de contraste et un antispasmodique au patient. Puis on introduit une sonde dans le rectum du patient afin d'y insuffler de l'air ou du dioxyde de carbone. L'acquisition des images s'effectue par des apnées de 25-30 secondes. L'acquisition de toutes les données s'effectue en environ 30 minutes.

L'examen ne requiert pas d'anesthésie et est donc moins invasif. La détection des polypes fonctionne parfaitement pour des polypes dont la taille est supérieure à 10 mm [ADW<sup>+</sup>13] (figure 1.12). Cependant, comme pour la capsule, il est impossible de retirer les lésions suspectes et on ne peut pas voir les lésions planes.

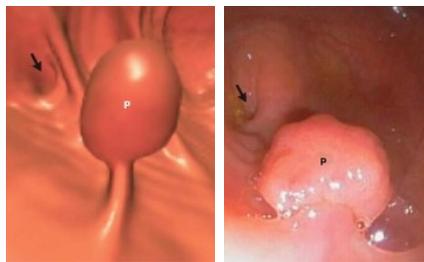


FIGURE 1.12 – Comparaison entre coloscopie virtuelle et la réalité

### 1.1.2 Les méthodes de détection assistée par ordinateur

Malgré toutes les méthodes matérielles (changement de capteurs, angle de vue, fausse couleur, contraste, etc.) mises au point pour améliorer le dépistage du cancer colorectal, et en particulier la détection des polypes, l'examen de référence reste la classique coloscopie (sauf pour des cas de contre indication à sa pratique). C'est pour cela que des méthodes d'aide à la détection assistée par ordinateur (CADe, Computer Aided Detection) sont apparues en ce qui concerne la détection des polypes colorectaux. De façon générale, les systèmes d'aide à la détection numérique sont des systèmes qui assistent les médecins pour l'interprétation des images médicales. Ces systèmes analysent donc les images des examens médicaux dans le but de mettre en évidence des zones suspectes, comme de possibles lésions, dans le but d'offrir au médecin un support pour sa décision. L'application la plus connue de ces systèmes est la détection des tumeurs, que ce soit pour la détection du cancer du sein [FTC<sup>+</sup>07], du cancer des poumons [MGH<sup>+</sup>16] ou du cancer colorectal.

Concernant la détection des polypes, que ce soit au cours des coloscopies ou que ce soit après l'examen (à partir de sa version vidéo enregistrée), il existe plusieurs méthodes afin de détecter les polypes avec des résultats intéressants. Elles s'organisent en trois grandes familles, qui sont les suivantes :

- Les méthodes ad-hoc (dites "hand-crafted")

- Les méthodes basées sur de l'apprentissage profond
- Les méthodes hybrides

Par la suite, ces méthodes vont donc être détaillées afin de décrire le paysage existant en termes d'outils d'aide à la détection dédiés aux polypes ou aux pathologies colorectales, et cela, en les rassemblant par famille de méthodes. Un tableau est proposé à la suite synthétisant les performances de chacune des méthodes présentées.

### Les méthodes ad-hoc

Les méthodes ad-hoc (ou "hand-crafted") sont formées à partir de différents blocs calculatoires, combinés ensemble, et essentiellement basés sur le traitement d'images. Les auteurs de ces méthodes se sont donc confrontés à la détection de pathologies colorectales, et en particuliers les polypes colorectaux. En voici donc un panel représentatif :

- Bernal et al. [BSFE<sup>+</sup>15] (figure 1.13) ont mis au point une méthode basée sur un modèle d'apparence des polypes qui définit les contours du polypes par rapport aux informations des vallées. C'est grâce à l'intégration des vallées qu'il leur est possible de générer des cartes d'énergie (WM-DOVA [BSV13]) qui relate la probabilité de présence d'un polype dans la zone en question. Les performances de leur méthode sont bonnes, et en particulier pour la détection de petits polypes, souvent ratés lors des coloscopies par les médecins. La figure 1.14 illustre les résultats de cette méthode.



FIGURE 1.13 – Illustration de la méthode de Bernal et al. [BSFE<sup>+</sup>15]

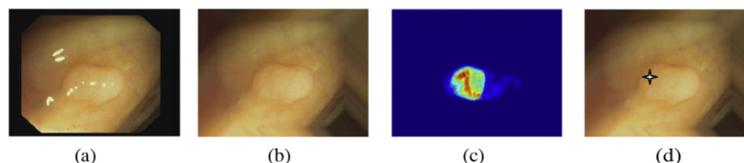


FIGURE 1.14 – Résultat de la méthode de Bernal et al. [BSFE<sup>+</sup>15], (a) image originale, (b) image prétraitée, (c) carte d'énergie, (d) probabilité de localisation du polype

- Zhu et al. [ZFL10] (figure 1.15) proposent une méthode, parfaitement adaptée à la coloscopie virtuelle, basée sur l'estimation de la courbure dans l'image puis ils appliquent la méthode de cartographie de Knutsson afin de détecter les polypes. Leur méthode permet d'améliorer l'estimation de la courbure, et donc permet de détecter des objets coloniques en plus des polypes comme les haustrations (il s'agit des plis de la paroi du côlon).

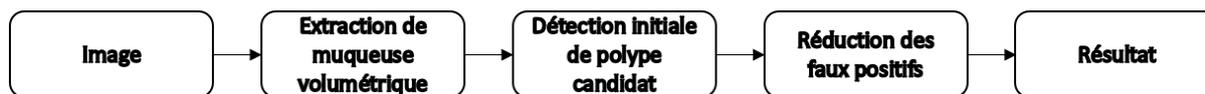


FIGURE 1.15 – Illustration de la méthode de Zhu et al. [ZFL10]

- Kang et al. [KD03] (figure 1.16) ont mis au point un système permettant de traiter les images endoscopiques en temps réel. Cette méthode a été développée dans le but de proposer un système rapide pour le traitement des images. Les images endoscopiques sont ainsi capturées et envoyées directement à des modules MATLAB. Ces modules permettent d’améliorer le contraste pour la détection, de segmenter l’image grâce à un filtre de détection de contours (Canny), puis pour finir, d’identifier les structures détectées. Cette identification est obtenue en comparant les paramètres détectés avec ceux caractéristiques d’un polype (pixel, segment, couleur, forme...). Cependant, aucune information sur les performances de cette méthode n’est donnée, si ce n’est qu’elle opère en temps réel.

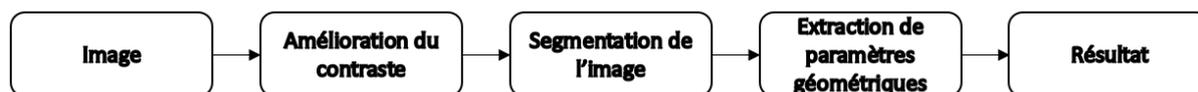
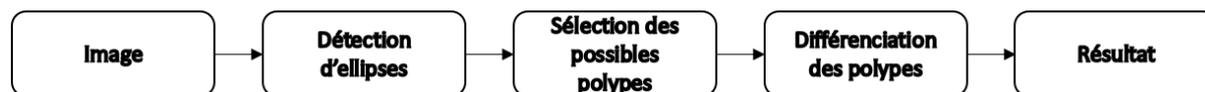


FIGURE 1.16 – Illustration de la méthode de Kang et al. [KD03]

- Hwang et al. [HOT<sup>+</sup>07] (figure 1.17) ont développé une méthode de détection des polypes basée sur la forme. La forme recherchée par Hwang et al. est celle d’une ellipse et se consacrent donc principalement à la détection d’ellipse dans les images endoscopiques. Pour y parvenir, ils segmentent l’image en plusieurs régions. Ils utilisent alors un algorithme de ligne de partage des eaux contrôlé par un marqueur pour la segmentation des polypes. Puis en utilisant les bordures des régions segmentées, ils génèrent l’ellipse correspondante. Enfin, pour sélectionner la zone correspondante à un polype parmi les zones détectées, plusieurs filtrages sont effectués (un filtrage par courbure et par direction de courbe, un filtrage par distance à une bordure, et un filtrage par intensité de pixel). Cette méthode obtient de bons résultats en terme de détection, mais au niveau du temps de calcul, elle n’est capable que de fonctionner à 15 images par secondes. La figure 1.18 montre des résultats de la détection.

FIGURE 1.17 – Illustration de la méthode de Hwang et al. [HOT<sup>+</sup>07]

- Tajbakhsh et al. [TGL16] (figure 1.19) proposent un moyen permettant de faire la détection automatique des polypes en coloscopie en utilisant les informations de formes et de contexte. L’information de contexte permet de réduire les structures

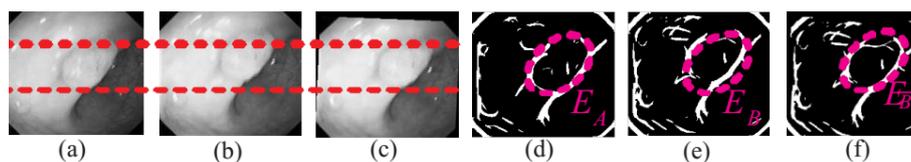


FIGURE 1.18 – Résultat de la méthode de Hwang et al. [HOT<sup>+</sup>07], (a) image en cours, (b) image suivante, (c) image précédente, (d) carte binaire de (a), (e) carte binaire de (b), (f) carte binaire de (c), avec la localisation du polype

ne contenant pas de polypes afin de localiser de manière fiable les polypes. Pour y arriver, ils obtiennent une carte brute des contours, puis ils retirent les contours des zones sans polypes, et finissent par localiser le polype grâce à des scores de probabilités sur la carte des contours. Cette méthode permet ainsi d’obtenir un taux de détection de 88% des polypes sur des images issues de coloscopies. La figure 1.20 représente des bonnes détections de polypes dans des images de coloscopies

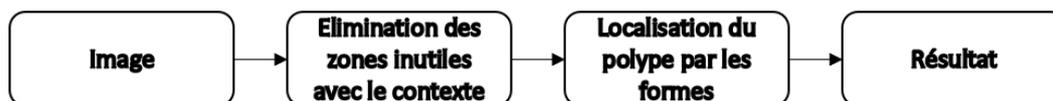


FIGURE 1.19 – Illustration de la méthode de Tajbakhsh et al. [TGL16]

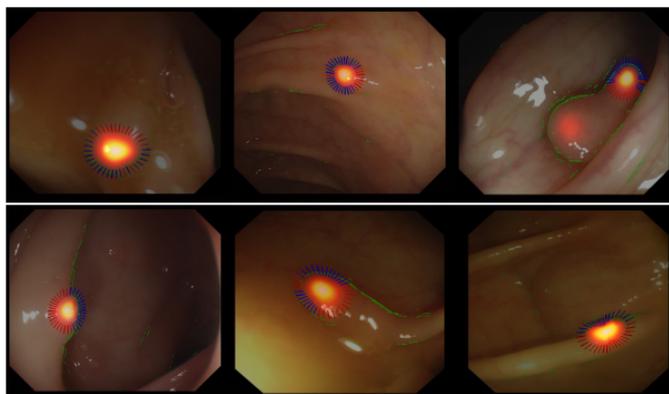


FIGURE 1.20 – Résultat de la méthode de Tajbakhsh et al. [TGL16]

- Sevo et al. [ŠAB<sup>+</sup>16] (figure 1.21) proposent de détecter les inflammations dans les vidéos d’examen coloscopiques en se basant sur la densité de bord. Ceci est basé sur une analyse de la texture. Une texture caractéristique d’un tissu enflammé est alors utilisée pour distinguer les tissus inflammatoires des tissus sains, en créant un filtre adapté. De plus, une version améliorée de cette méthode permet aussi d’éliminer les vaisseaux sanguins. Les résultats de cette méthode montrent que les régions inflammatoires peuvent être détectées en temps réel avec une précision supérieure à 84%.

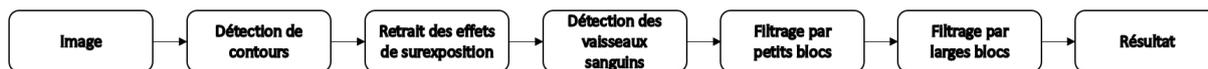


FIGURE 1.21 – Illustration de la méthode de Sevo et al. [ŠAB+16]

### Les méthodes basées sur de l'apprentissage profond

La seconde famille de méthodes est celle des méthodes basées sur l'apprentissage profond. Elle se consacre aux méthodes qui n'utilisent que l'apprentissage machine profond dans leur chaîne de traitement afin d'effectuer la détection. Dans la littérature, de nombreuses méthodes sont proposées, en voici ici les plus représentatives :

- Park et al. [PS16] (figure 1.22) proposent de détecter les polypes colorectaux en utilisant un réseau de neurones convolutif. Ils ont appliqué un modèle d'image propre pour extraire les caractéristiques représentant des polypes, des tissus normaux, et bien d'autres structures à partir de vidéos d'examen. La classification a été effectuée en utilisant un modèle de champ aléatoire conditionnel. Puis, ils remplacent les descripteurs de caractéristiques d'image propre avec des fonctionnalités d'un réseau de neurones convolutif. L'utilisation du réseau de neurones montre une plus grande robustesse concernant la prise de vue autour du polype. De plus, ils obtiennent de bons résultats en terme de sensibilité et de spécificité (plus de 85%). La figure 1.23 montre le résultat d'une détection sur une image.

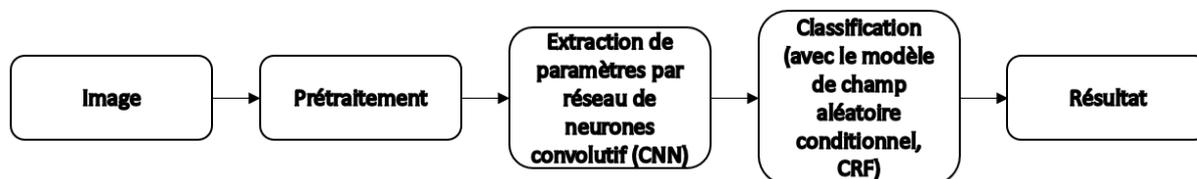


FIGURE 1.22 – Illustration de la méthode de Park et al. [PS16]

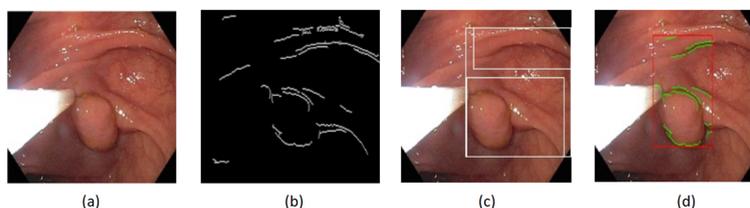


FIGURE 1.23 – Résultat de la méthode de Park et al. [PS16], (a) image de vidéo, (b) détection des contours, (c) projection de fenêtres, (d) résultat de la classification CRF

- Ribeiro et al. [RUH16] (figure 1.24), eux aussi, utilisent les réseaux de neurones convolutifs pour effectuer la classification des polypes colorectaux. Pour y parvenir, ils se basent sur l'extraction de paramètres de texture de l'image afin d'en effectuer la classification. Afin de rendre la classification d'autant plus efficace, ils décident d'utiliser non pas les images normales pour effectuer l'apprentissage, mais

des patches afin d'augmenter la quantité de données pour apprendre. Leur méthode permet d'atteindre une précision de 90% concernant la classification des polypes.

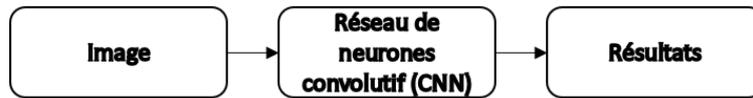


FIGURE 1.24 – Illustration de la méthode de Ribeiro et al. [RUH16]

### Les méthodes hybrides

Puis pour finir, il existe les méthodes hybrides, ces méthodes combinent principalement différents paramètres (texture, forme, contexte...) pour détecter des régions d'intérêt avec de l'apprentissage machine afin de discriminer ces régions d'intérêt et de trouver la pathologie recherchée. Voici quelques unes des méthodes les plus représentatives qui existent à ce jour :

- Iwahori et al. [ISH<sup>+</sup>13] (figure 1.25) proposent une technique de détection des polypes en utilisant un filtre Hessian et les techniques d'apprentissage machine. Dans leur technique, le filtre Hessian est utilisé pour détecter les régions dans les images qui auraient une forme de "blob" (la forme de certains polypes). Puis une extraction des paramètres de couleurs et de contours s'en suit pour ces régions. Finalement, une machine à vecteurs de support (SVM, Support Vector Machine) est utilisée sur ces paramètres pour déterminer les possibles régions contenant un polype. Cette méthode permet d'obtenir une grande précision expérimentalement. La figure 1.26 montre le résultat de cette méthode sur une image.



FIGURE 1.25 – Illustration de la méthode d'Iwahori et al. [ISH<sup>+</sup>13]

- Silva et al. [SHR<sup>+</sup>14] (figure 1.27) ont développé une méthode permettant de détecter les polypes dans les images de capsules endoscopiques en combinant l'extraction de paramètres géométriques dans l'image et l'apprentissage machine. La détection des formes permet de créer des régions d'intérêt dans l'image. Ainsi, des paramètres de textures sont alors générés à partir de ces régions et une comparaison de ces paramètres par rapport à des résultats issus de l'apprentissage machine permet de déterminer s'il s'agit d'un polype ou non. Les résultats montrent de bons résultats en terme de sensibilité et de spécificité.

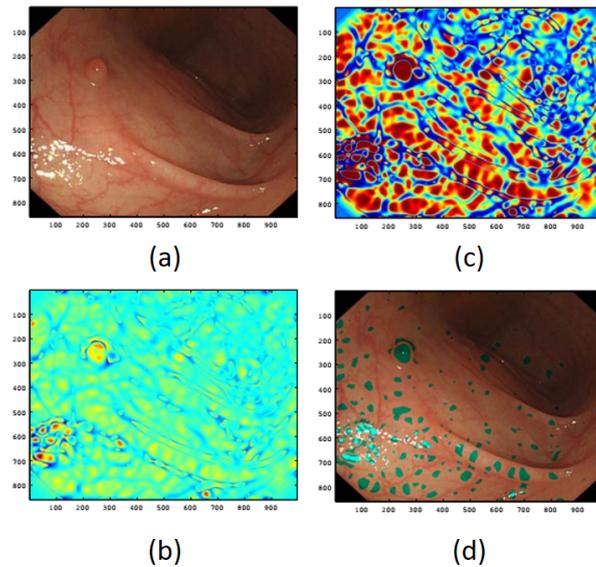


FIGURE 1.26 – Résultat de la méthode d’Iwahori et al. [ISH<sup>+</sup>13], (a) image originale, (b) image améliorée, (c) image segmentée, (d) régions candidates

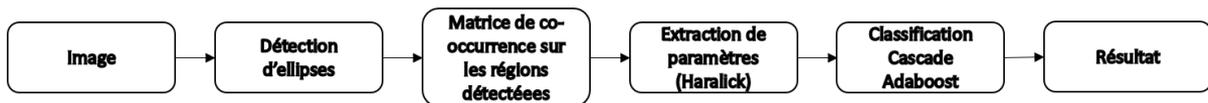


FIGURE 1.27 – Illustration de la méthode de Silva et al. [SHR<sup>+</sup>14]

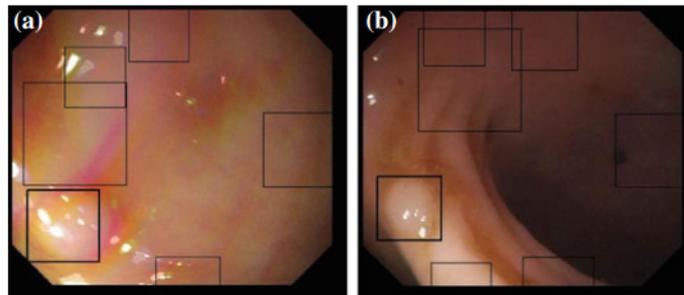
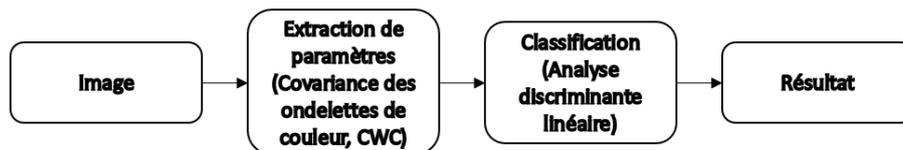
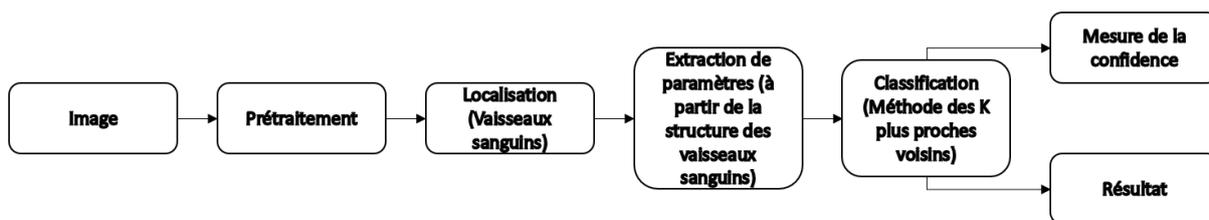


FIGURE 1.28 – Résultat de la méthode de Silva et al. [SHR<sup>+</sup>14], les rectangles noirs sont toutes les zones détectées, les rectangles noirs plus épais sont les zones classées comme possible polype

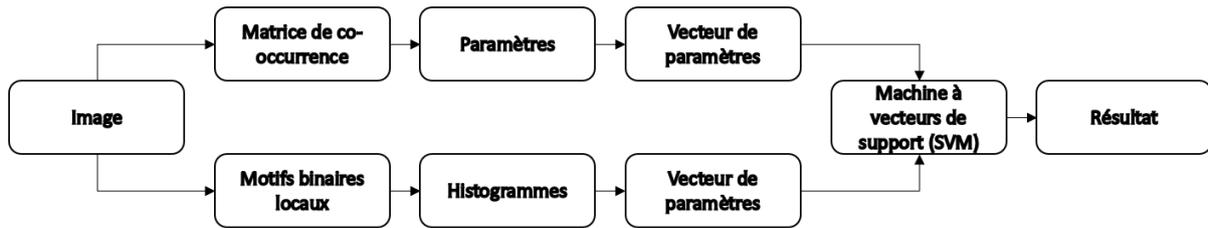
- Karkanis et al. [KIM<sup>+</sup>03] (figure 1.29) ont mis au point un outil d’aide à la détection dans les vidéos utilisant les caractéristiques des ondelettes de couleur. Ces caractéristiques sont basées sur la mesure des textures du second ordre de la covariance. Cet outil est basé sur une analyse discriminante linéaire (LDA, Linear Discriminant Analysis), méthode comparable aux méthodes d’apprentissages supervisées, afin de caractériser les différentes régions d’une image afin de détecter le polype. Cet outil propose de bons résultats en terme de spécificité et sensibilité.
- Gross et al. [GSB<sup>+</sup>09] (figure 1.30) ont développé une méthode effectuant la classifi-

FIGURE 1.29 – Illustration de la méthode de Karkanis et al. [KIM<sup>+</sup>03]

cation des polypes en se basant sur des paramètres extraits des vaisseaux sanguins. Ils proposent alors un moyen permettant de faire la différence entre un polype hyperplasique et un polype adénomateux. Pour y parvenir, une classification est effectuée sur l'extraction des paramètres relatant la structure des vaisseaux sanguins. Ils combinent la longueur de ces vaisseaux, leur périmètre moyen et la valeur de l'intensité moyen (en niveaux de gris) de ces vaisseaux. Puis en comparant les résultats de leur méthode par rapport aux LBP (Local Binary Pattern, motif binaire local en français), ils montrent que leur solution est capable de bien classer les polypes dans 90% des cas.

FIGURE 1.30 – Illustration de la méthode de Gross et al. [GSB<sup>+</sup>09]

- Ameling et al. [AWP<sup>+</sup>09] (figure 1.31), proposent une méthode de détection des polypes basée sur la texture. Analyser la texture passe par l'extraction de paramètres dédiés, en l'occurrence ici les paramètres d'Haralick issus de la matrice de cooccurrence en niveaux de gris (GLCM, Grey Level Cooccurrence Matrix) et les motifs binaires locaux (LBP, Local Binary Pattern). Et pour venir suppléer l'extraction de ces paramètres et ainsi détecter les polypes, les auteurs utilisent des machines à vecteurs de support (SVM). Un modèle de polype est ainsi obtenu à partir de données d'entraînement. Les résultats obtenus montrent principalement que l'analyse de la texture des images permettent d'obtenir de bonnes détections des polypes.
- Bae et al. [BY15] (figure 1.32) ont mis au point une méthode permettant de détecter les polypes en utilisant l'apprentissage déséquilibré et l'apprentissage discriminatif. Ils se sont proposés d'utiliser l'apprentissage déséquilibré en raison du déséquilibre qu'il existe dans les bases de données de polypes (entre zones positives et zones négatives). Afin d'y parvenir, dans leur processus d'apprentissage, ils créent des classificateurs faibles avec des jeux de données rééquilibrés par échantillonnage ascendant/descendant, et de générer un classificateur fort en combinant les faibles. Puis pour améliorer la discrimination entre polype et non polype, ils proposent une méthode d'apprentissage d'analyse des moindres carrés partielle. Les résultats

FIGURE 1.31 – Illustration de la méthode d’Ameling et al. [AWP<sup>+</sup>09]

montrent alors de nettes améliorations par rapport aux détecteurs classiques. La figure 1.33 montre des résultats de détections.

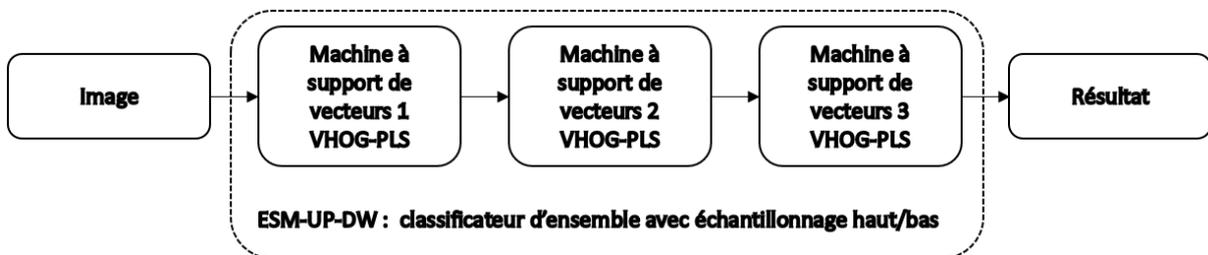


FIGURE 1.32 – Illustration de la méthode de Bae et al. [BY15]

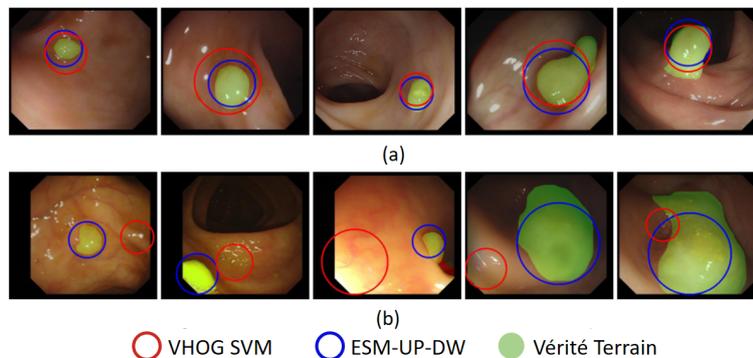


FIGURE 1.33 – Résultat de la méthode de Bae et al. [BY15], (a) cas de bonnes détections, (b) cas de mauvaises détections

Le tableau 1.1 permet de comparer les performances des différentes méthodes proposées ici. Ce tableau a pour but de montrer qu’il n’existe pas à ce jour de méthode unique pour mesurer les performances d’une approche permettant d’aider à la détection des polypes. On peut ainsi constater des informations complètement disparates. Premièrement, les données (images) utilisées pour tester les méthodes sont totalement différentes, on ne peut alors pas vraiment se rendre compte des performances d’une méthode par rapport à une autre. Deuxièmement, les métriques utilisées pour mesurer ces performances sont tout aussi différentes, il peut être alors difficile de juger si telle ou telle méthode est meilleure par rapport à une autre. Par ailleurs, aucune de ces méthodes n’a, aujourd’hui, été trans-

férée en routine clinique.

Ainsi, à la vue des résultats proposés, et afin de détecter les polypes colorectaux, il peut être intéressant de proposer une méthode se plaçant dans la famille des méthodes hybrides afin de tirer parti des avantages de l'extraction des paramètres de l'image combinée à une étape d'apprentissage machine afin de réduire le nombre de fausses détections et ainsi obtenir une méthode performante. Mais plus encore, la contrainte temps réel doit être présente dans la méthode que l'on souhaite proposer. En effet, dans l'optique de proposer un outil de détection des polypes colorectaux qui soit facilement intégrable et permettant d'assister le médecin durant l'examen et non pas à posteriori (chose qui induit un second examen si le médecin s'aperçoit qu'il a raté un polype), on se doit de proposer un outil fonctionnant sans perturber la routine clinique.

Un autre objectif apparait ici. En effet, comme il n'existe pas d'uniformité dans le protocole de mesures des performances, il serait intéressant de proposer des métriques de référence qui permettraient de mieux rendre compte des performances des outils proposés, sur des bases de données ouvertes, clairement identifiées.

| Auteurs                               | Méthode | Données                        | Performances   | Temps de calcul  |
|---------------------------------------|---------|--------------------------------|--|------------------|
| Tajbakhsh et al. [TGL16]              | AH      | CVC-ColonDB                    | Sensibilité : 88%  | 2.6 s/image      |
| Sevo et al. [ŠAB <sup>+</sup> 16]     | AH      | 3 vidéos (6375 images)         | Exactitude > 84%   | 13 ms/image      |
| Bernal et al. [BSFE <sup>+</sup> 15]  | AH      | CVC-ClinicDB (612 images)      | Exactitude : 70.26%  | X                |
| Zhu et al. [ZFL10]                    | AH      | 52 sets d'images               | Sensibilité : 100%, 3.5 fausses détections par set           | X                |
| Kang et al. [KD03]                    | AH      | X                              | X  | 1 image/s        |
| Hwang et al. [HOT <sup>+</sup> 07]    | AH      | 8621 images                    | Rappel : 96%, Précision : 84%                                | 15 images/s      |
| Park et al. [PS16]                    | AP      | 35 vidéos                      | Sensibilité : 86%, Spécificité : 85%                         | X                |
| Ribeiro et al. [RUH16]                | AP      | 100 images                     | Sensibilité : 95.16%, Spécificité : 74.19%                   | X                |
| Ameling et al. [AWP <sup>+</sup> 09]  | HY      | 4 heures de vidéo              | Aire sous la courbe : 0.95                                   | X                |
| Silva et al. [SHR <sup>+</sup> 14]    | HY      | 500 images                     | Sensibilité : 91%, Spécificité : 95%                         | 37.5ms/image     |
| Bae et al. [BY15]                     | HY      | CVC-ColonDB + 1263 images      | Rappel > 90%   | 0.6375 s/image   |
| Iwahori et al. [ISH <sup>+</sup> 13]  | HY      | 128 images                     | Exactitude : 96.8%, Spécificité : 98.2%, Sensibilité : 95.4% | X                |
| Gross et al. [GSB <sup>+</sup> 09]    | HY      | 102 images                     | Rappel : 90%   | X                |
| Karkanis et al. [KIM <sup>+</sup> 03] | HY      | 1200 images (20 images/vidéos) | Spécificité : 97%, Sensibilité : 90%                         | 1.6 minute/vidéo |

TABLE 1.1 – Tableau comparatif du panel exhaustif des méthodes existantes pour la détection des polypes colorectaux (AH = Ad-Hoc, AP = Apprentissage Profond, HY = HYbride)

## 1.2 Apprentissage machine orienté Boosting

L'apprentissage machine est donc une étape incontournable afin de mettre au point un outil pour effectuer la détection des polypes en coloscopie. Parmi toutes les méthodes d'apprentissage machine existantes, le boosting et quelques unes de ses variantes n'ont été que peu utilisées mais sont pourtant très prometteuses. En effet, le boosting est un bon compromis en termes de performances de détection, de temps de calcul et de capacité à être embarqué.

### 1.2.1 Boosting

Le boosting repose sur la combinaison de classificateurs [Sch90] [Fre95]. Il fonctionne par itérations successives, en partant de classificateurs faibles afin de construire un classificateur fort. On appelle souvent un apprenant faible un algorithme qui permet de construire un classificateur fort, dont les performances de classifications sont légèrement supérieures au hasard. Le classificateur final obtenu est une combinaison linéaire des classificateurs faibles précédents. Les exemples qui sont mal classés sont alors boostés afin d'avoir plus de poids dans la classification et donc plus d'importance par rapport à l'apprenant faible au prochain tour d'apprentissage, afin de pallier le manque.

Il existe de nombreuses méthodes de boosting, et parmi elles, la méthode de boosting adaptatif, communément appelé Adaboost. C'est grâce à cette méthode d'apprentissage que l'on va mettre au point notre outil afin d'en améliorer les performances et en particulier, la précision, tout en conservant un temps de traitement compatible avec la contrainte vidéo.

### 1.2.2 Adaboost

D'après [VJ01], l'algorithme d'apprentissage Adaboost est utilisé afin d'augmenter les performances de classification d'un algorithme d'apprentissage simple (on peut prendre l'exemple d'augmenter les performance d'un simple perceptron). Ceci est obtenu en combinant un ensemble de fonctions de classification faible afin de former un classificateur plus fort. Dans le langage du boosting, un algorithme d'apprentissage faible est appelé faible apprenant.

L'apprenant est appelé faible parce que personne ne s'attend à ce que la meilleure fonction de classification classe correctement les données d'apprentissage (c'est-à-dire pour un problème donné, le meilleur perceptron ne peut que classer les données de formation correctement que dans 51% des cas). Pour que l'apprenant faible soit stimulé, il est appelé à résoudre une séquence de problèmes d'apprentissage. Après le premier cycle d'apprentissage, les exemples sont redéfinis pour souligner ceux qui ont été mal classés par le classificateur faible précédent. Le classificateur final, dans notre exemple, prend la forme d'un perceptron construit comme une combinaison linéaire pondérée de classificateurs faibles suivie d'un seuil.

L'algorithme décrit dans le tableau 1 sert à sélectionner les classificateurs clés faibles à partir de l'ensemble des classificateurs faibles possibles. Pour que le processus AdaBoost soit assez efficace, l'ensemble de classificateurs faibles peut être fortement grand

- 
1. On considère les images d'apprentissage  $(x_1, y_1), \dots, (x_n, y_n)$  où  $y_i = 0$  pour les exemples négatifs (respectivement 1 pour les exemples positifs)
  2. On initialise les poids  $w_{1,i} = \frac{1}{2m}$  pour  $y_i = 0$  (respectivement  $w_{1,i} = \frac{1}{2l}$  pour  $y_i = 1$ ) avec  $m$  et  $l$  les nombres d'images négatives et positives
  3. Pour  $t = 1, \dots, T$  :
    - On normalise les poids  $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$
    - On choisit le plus faible classificateur grâce à l'erreur basée sur le poids :

$$\epsilon_t = \min_{f,p,\theta} \sum_i w_i |h(x_i, f, p, \theta) - y_i|$$

où  $h(x_i, f, p, \theta)$  correspond à un classificateur faible avec un paramètre ( $f$ ), un seuil ( $\theta$ ) et une polarité ( $p$ ) qui indique la direction de l'inégalité

- On définit  $h_t(x) = h(x, f_t, p_t, \theta_t)$  où  $f_t, p_t$ , et  $\theta_t$  sont les "minimizers" de  $\epsilon_t$
- On met à jour les poids :

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$

où  $e_i = 0$  si l'exemple  $x_i$  est classé correctement,  $e_i = 1$  sinon, et  $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

4. Le classificateur fort final est :

$$C(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{sinon} \end{cases}$$

où  $\alpha_t = \log\left(\frac{1}{\beta_t}\right)$

---

TABLE 1.2 – Algorithme de boosting.  $T$  hypothèses sont faites en utilisant chaque paramètre. L'hypothèse finale est une combinaison linéaire pondérée des  $T$  hypothèses où les poids sont inversement proportionnels à l'erreur d'apprentissage

### 1.2.3 Cascade Adaboost

D'après [VJ04], l'algorithme utilisant une cascade de classificateurs permet de réaliser de meilleures détections tout en réduisant le temps de calcul. L'essentiel dans cet algorithme est qu'il utilise des classificateurs plus efficaces qui rejettent une bonne partie des sous-fenêtres négatives tout en détectant la quasi totalité des zones positives dès les premiers étages de la cascade. Les classificateurs plus simples sont utilisés pour rejeter la majorité des sous-fenêtres avant que des classificateurs plus complexes ne soient appelés à obtenir des faibles taux de faux positifs.

Les étages de la cascade sont construits en entraînant des classificateurs avec l'apprentissage Adaboost.

La forme globale du processus de détection est celui d'un arbre de décision, qui est appelé "cascade" (voir figure 1.34) [Quinlan, 1986]. Un résultat positif du premier classificateur engage l'évaluation à travers le second classificateur qui a été aussi ajusté afin d'avoir un fort taux de détection. Un résultat positif du second classificateur va engager l'évaluation à travers le troisième classificateur, et ainsi de suite. Cependant, un résultat négatif à la sortie de n'importe lequel des classificateurs va entraîner le retrait immédiat de la cascade de la sous-fenêtre traitée. La structure de la cascade reflète donc le fait que parmi n'importe quelle image, une majorité de sous-fenêtres sont négatives. La cascade tente de rejeter autant d'éléments négatifs que possible et en le faisant le plus tôt possible dans les étages de la cascade. Bien qu'une instance positive déclenche l'évaluation de chaque classificateur dans la cascade, il s'agit d'un événement extrêmement rare. Plus comme les arbres de décisions, les classificateurs suivants sont entraînés en utilisant les exemples qui ont réussi à passer à travers les étages précédents. Il en découle alors que le second classificateur fait face à une tâche plus difficile que le premier classificateur. L'exemple qui arrive à passer le premier étage est théoriquement plus compliqué à déterminer que des exemples typiques.

En pratique, un framework simple est utilisé pour construire un classificateur à la fois efficace et efficient. L'utilisateur doit sélectionner le taux maximum acceptable de fausses alarmes par étage et le taux minimum de détection acceptable par étage. Chaque étage de la cascade est entraînée en utilisant la technique Adaboost dont le nombre de paramètres utilisés augmente tant que les taux cibles de fausses alarmes et de détection acceptable sont atteints pour l'étage en cours d'apprentissage. Ces taux sont déterminés en testant le détecteur courant sur un set de données de validation. Si le taux global de fausses détections n'est toujours pas atteint, alors un nouvel étage est ajouté à la cascade. Le set de données d'exemples négatifs pour entraîner les étages suivants est obtenu en collectant toutes les fausses détections en testant le détecteur courant sur un set d'image qui ne contient aucune instance des choses à détecter. Ce framework est disponible dans le tableau 1.3. L'application la plus connue de ce framework est la détection des visages dans les images, qui est illustré par la figure 1.35.

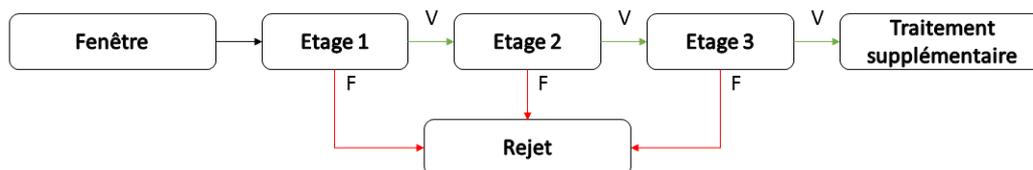


FIGURE 1.34 – Illustration du principe de détection Cascade Adaboost

### 1.3 Conclusion

On a pu voir dans cette partie les moyens mis en place (matériels ou logiciels) permettant d'aider le médecin dans l'amélioration de ses capacités à mieux détecter les polypes

- 
1. L'utilisateur choisit des valeurs pour  $f$  (Le taux maximum acceptable de fausses alarmes par étage) et  $d$  (le taux minimum de détection acceptable par étage)
  2. L'utilisateur choisit aussi le taux cible de fausses détections total  $F_{target}$
  3.  $P$  = set d'exemples positifs
  4.  $N$  = set d'exemples négatifs
  5.  $F_0 = 1.0$ ,  $D_0 = 1.0$
  6.  $i = 0$
  7. Tant que  $F_i > F_{target}$ 
    - $i \leftarrow i + 1$
    - $n_i = 0$ ,  $F_i = F_{i-1}$
    - Tant que  $F_i > f \times F_{i-1}$ 
      - $n_i \leftarrow n_i + 1$
      - On utilise  $P$  et  $N$  pour entrainer un classificateur avec  $n_i$  paramètres en utilisant l'Adaboost
      - On évolue le classificateur cascade actuel avec un set de validation pour déterminer  $F_i$  et  $D_i$
      - On baisse le seuil pour le  $i^{\text{ème}}$  classificateur tant que le classificateur cascade actuel a un taux de détection d'au moins  $d \times D_{i-1}$  (qui affecte aussi  $F_i$ )
    - $N \leftarrow \emptyset$
    - Si  $F_i > F_{target}$  alors on évalue le détecteur cascade actuel sur un set d'images négatives et en mettant toutes les fausses détections dans le set  $N$
- 

TABLE 1.3 – Framework d'apprentissage Cascade Adaboost pour obtenir un détecteur cascade

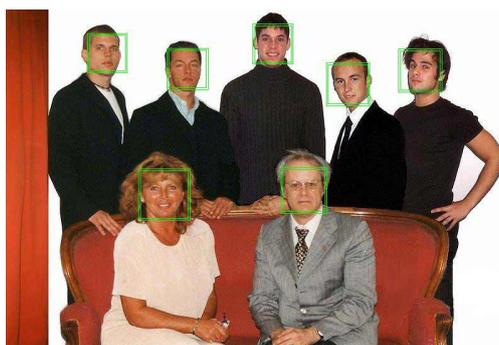


FIGURE 1.35 – Illustration du résultat de la détection Cascade Adaboost (Source : OpenCV Docs)

lors d'un examen de coloscopie. Malgré les solutions matérielles proposées, on a pu s'apercevoir que la coloscopie classique reste l'examen de référence et que peu, voire aucune, de ces optimisations ne sont utilisées actuellement en routine clinique. Cela a donc donné naissance à de nouveaux outils d'aide à la détection assistée par ordinateur, qui ont pour

but de s'adapter à la coloscopie classique. C'est donc pour cela qu'il peut être intéressant de proposer un tel outil qui sera susceptible d'intéresser les médecins s'il propose des performances acceptables tout en ne modifiant pas le déroulé de l'examen (on pourrait ici parlé d'une méthode user friendly).

Si on souhaite mettre au point un outil permettant d'aider les médecins à réduire le taux de polypes manqués, il faut alors proposer un système d'aide à la détection assistée par ordinateur, car il s'agit d'un outil que les praticiens pourront utiliser plus facilement dans le cadre de l'examen. De plus, parmi les 3 familles de méthodes, seule la famille des méthodes hybrides est la plus adaptée. En effet, cette famille propose de bonnes performances de détection associées à des temps de calcul pouvant s'approcher du temps réel. Cela répond en effet à plusieurs contraintes que l'on s'est fixées, qui sont l'aspect temps réel et les bonnes performances de détection.

Par la suite, il est proposé de mettre au point des méthodes hybrides qui vont permettre de détecter les polypes dans les images issues de coloscopies. L'utilisation de l'apprentissage machine, et précisément l'utilisation de la méthode de Viola et Jones, méthode dont les performances en termes de temps de calcul sont plus qu'intéressantes, va être plus qu'incontournable dans la mise au point de notre outil.

# Chapitre 2

## La détection des polypes dans les images issues de coloscopies

Dans ce chapitre, on va s'intéresser à la détection des polypes issus d'images fixes. Et pour y parvenir, deux méthodes mises au point afin de détecter automatiquement les polypes dans les images de coloscopies sont introduites.

Les deux méthodes introduites ici afin de détecter les polypes dans les images de coloscopies sont toutes les deux des méthodes hybrides qui vont combiner extraction de paramètres et apprentissage machine afin d'obtenir des bonnes performances de détection et des temps de calcul compatibles avec les 25 images par seconde d'un coloscope clinique (soit 40 millisecondes).

Pour rappel, on cherche à mettre au point un outil répondant aux critères suivants :

- être capable de détecter les polypes en temps réel (c'est-à-dire que le temps pour traiter une image issue de coloscopie doit être inférieur à 40 millisecondes)
- être compatible avec du matériel dont les ressources de calculs sont standards et limitées
- obtenir des performances élevées en termes de sensibilité et de spécificité
- être compatible avec l'utilisation en routine clinique (et donc ne pas modifier la pratique clinique du médecin) et alerter le médecin sur la présence de polypes durant l'examen.

Pour les deux méthodes présentées, on va s'intéresser à leur fonctionnement mais aussi leurs performances et limitations, dans le but de déterminer laquelle est la plus adaptée à notre problématique de fournir un outil capable de détecter une grande majorité de polypes. Et dans un souci d'aller vers une application clinique de cet outil, on cherchera à savoir laquelle des deux méthodes est la plus prometteuse afin de détecter les polypes en temps réel.

## 2.1 Une méthode pour la détection des polypes basée sur la forme

### 2.1.1 Constat sur la forme des polypes

Quand on regarde des images contenant des polypes (figure 2.1) extraites d'examens coloscopiques, on peut effectuer le constat suivant : la forme des polypes est principalement soit circulaire, soit elliptique. Ainsi, on peut donc déduire qu'il y a un intérêt tout trouvé à proposer une méthode permettant de détecter ces formes dans ces images dans l'objectif de réduire le taux de polypes manqués durant l'examen. Si on recherche directement les formes circulaires dans les images, cela permet de réduire les étapes de traitement qui peuvent suivre en limitant le nombre de région d'intérêt et de ne pas traiter la totalité de l'image. Cette méthode se doit d'être embarquable pour être utilisée dans les salles d'examens, et donc de fonctionner en temps réel pour que l'information de la détection soit disponible le plus rapidement possible afin d'en informer le médecin. On peut alors proposer la chaîne de traitement disponible en figure 2.2 pour la détection des polypes basée sur les formes circulaires.

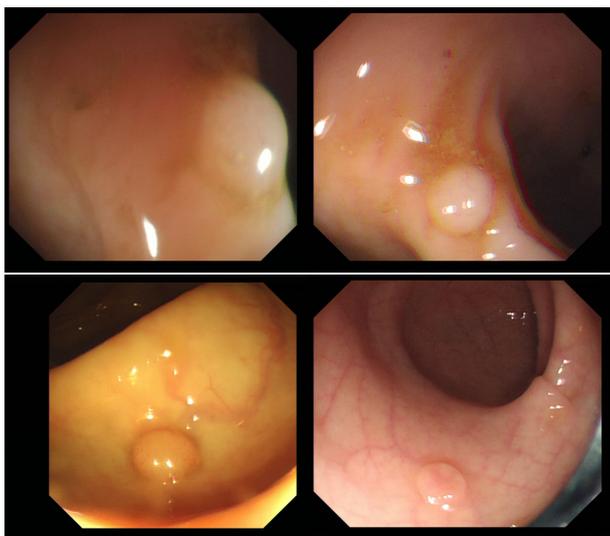


FIGURE 2.1 – Illustration concernant la forme des polypes

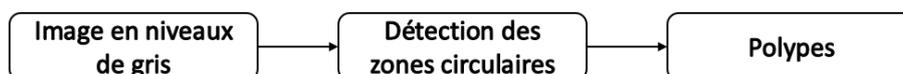


FIGURE 2.2 – Chaîne de traitement pour la détection des polypes basée sur la forme

### 2.1.2 La détection des formes circulaires

Il existe de nombreux moyens pour détecter les formes circulaires dans les images avec plus ou moins d'avantages et d'inconvénients. Par exemple, on peut citer les méthodes suivantes : la transformée de Hough circulaire [Dav88], la méthode Log-Gabor, la méthode RANSAC [Der10] et bien d'autres qui sont disponibles dans la littérature. Cependant, le choix ici s'est porté sur deux méthodes en particulier : la transformée de Hough circulaire et la méthode Log-Gabor.

Ces deux méthodes vont être introduites par la suite et seront comparées afin de montrer les avantages et inconvénients de chacune de ces deux méthodes puis de choisir la méthode la plus adaptée à la détection des polypes dans les images de coloscopies.

#### La transformée de Hough

La transformée de Hough est une méthode introduite par Hough [Hou62] en 1962. Elle permet de retrouver l'ensemble des lignes présentes dans une image quelconque. Le principe de cette transformée est qu'elle suppose qu'il existe un nombre infini de lignes qui peuvent passer par un point de contours, et dont la seule différence notable est l'angle d'orientation. La finalité de cette méthode est de déterminer quelles sont les lignes qui sont les plus fréquentes dans l'image analysée. La figure 2.3 illustre le résultat de la transformée de Hough sur une image afin d'en extraire les lignes droites. Afin d'y parvenir, une détection des contours est effectuée dans un premier temps, puis la transformée de Hough est appliquée.

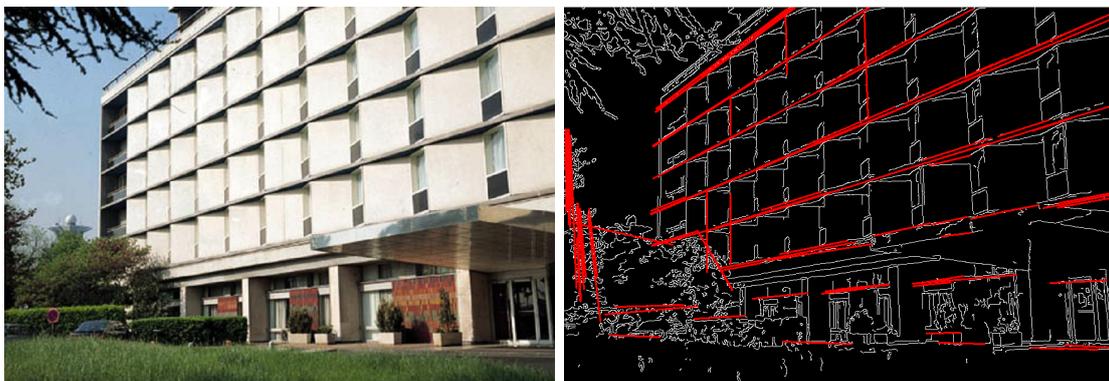


FIGURE 2.3 – Illustration de la transformée de Hough (Source : OpenCV Docs)

Cependant, la méthode proposée par Hough [Hou62] permet de détecter uniquement les lignes mais Duda [DH72] propose un moyen d'adapter la transformée de Hough pour détecter des cercles dans une image, et de façon générale, les courbes dans une image. Il suffit alors de changer de représentation dans l'image. Pour les cercles, il suffit de choisir une représentation paramétrique pour la famille de tous les cercles et de transformer chaque point de la figure de la bonne manière. Selon [Dav88], on envisage alors toutes les

positions de centres possibles, à une distance  $\mathcal{R}$  de chaque pixel unique, qui sont accumulées dans l'espace paramétrique,  $\mathcal{R}$  étant le rayon du cercle cherché. L'espace paramétrique peut être considérée comme une zone de stockage, mais lorsqu'il s'agit de rechercher un cercle, il doit être congruent à l'espace de l'image : les centres de cercles possibles sont alors accumulés dans un nouveau plan de l'espace de l'image. Enfin, les pics dans l'espace paramétrique correspondent aux centres d'objets circulaires. La figure 2.4 donne une idée de comment est obtenu le centre d'un objet circulaire.

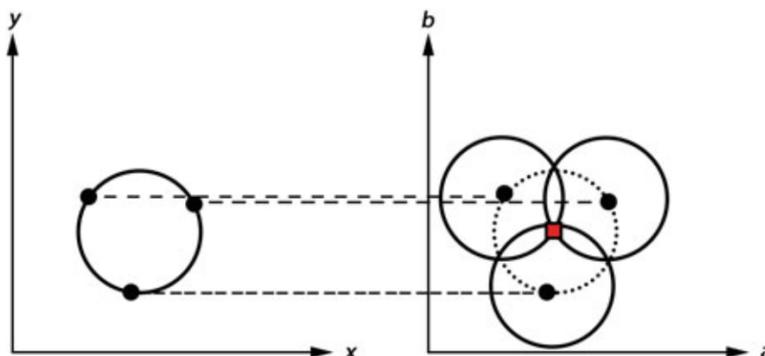


FIGURE 2.4 – Illustration de la transformée de Hough circulaire. Sur la gauche, trois motifs circulaires dans l'espace de l'image, et sur la droite, trois cercles correspondants au point rouge dans l'espace paramétrique (de coordonnées  $(a, b)$  pour un rayon  $\mathcal{R}$  fixé)

Côté ressources de calcul, le processus afin d'appliquer la transformée de Hough reste simple et efficient [LLLM86][ZIN14]. Ce qui est un avantage dans l'optique d'embarquabilité de la méthode et donc d'un fonctionnement dans un environnement disposant de ressources de calcul limitées.

### La méthode des filtres Log-Gabor

Karagyris et al. présente dans son article [KB09] dédié à la vidéocapsule une méthode basée sur les filtres Log-Gabor. Les filtres Log-Gabor sont basés sur les filtres de Gabor. Les filtres Log-Gabor ont une réponse gaussienne sur une échelle de fréquence logarithmique au lieu de linéaire comme les filtres de Gabor. Ils permettent alors d'avoir une bande passante optimisée afin de produire un filtre avec une étendue spatiale limitée [Kov00]. L'avantage le plus important des filtres Log-Gabor est que, par leur queue étendue dans les hautes fréquences, ils permettent d'encoder de façon plus efficace les images naturelles. Globalement, le filtre Log-Gabor ressemble donc au système visuel humain.

Combiné au filtres Log-Gabor, Karagyris et al. utilisent en complément le détecteur de contour SUSAN (pour Smallest Univalued Segment Assimilating Nucleus) décrit dans [SB97]. Le but est d'améliorer les contours du résultats du filtrage par la méthode des filtres Log-Gabor. Et pour compléter cela afin de détecter les polypes dans les images de coloscopies, un algorithme parcourt les contours de l'image binaire obtenue par les différents filtrages afin de calculer la courbure de chaque point. Pour chacun de ces points,

## 2.1. UNE MÉTHODE POUR LA DÉTECTION DES POLYPES BASÉE SUR LA FORME<sup>49</sup>

les centres de courbures sont trouvés et stockés. Cela va résulter en un nuage de points à l'intérieur de l'image binaire. Puis un simple seuillage permet de créer des régions simples. Le résultat de cette méthode est observable sur les figures 2.5 et 2.6. Enfin, un contrôle de chacune des régions trouvées permet de déterminer selon 2 critères si ces régions sont possiblement des polypes ou non.

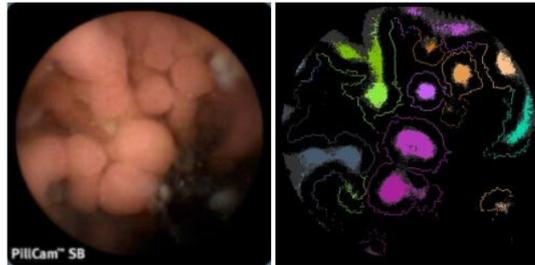


FIGURE 2.5 – Résultat de la détection basée sur les filtres Log-Gabor [KB09]

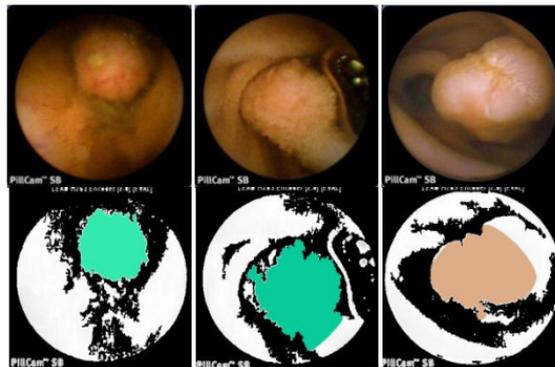


FIGURE 2.6 – Résultat de la détection basée sur les filtres Log-Gabor [KB09]

### Conclusion intermédiaire sur la méthode à adopter

Les deux méthodes proposent donc toutes les deux des intérêts. Mais le critère qui va permettre de les départager est de tester leurs performances en terme de détection. Pour cela, on décide de tester ces deux méthodes sur l'ensemble des images disponibles dans la base de données CVC-ColonDB [BSV12]. Afin de quantifier les performances, on utilise juste trois métriques parmi les suivantes (en gras), et les différents résultats sont disponibles dans le tableau 2.1.

- Les vrais positifs (True Positive, TP). Une détection "vraie positive" est le résultat lorsqu'un polype est détecté correctement.
- Les faux positifs (False Positive, FP). Une détection "fausse positive" est le résultat lorsqu'un polype est signalé négatif alors qu'il devrait être positif.

- Les faux négatifs (False Negative, FN). Une détection "fausse négative" est le résultat lorsqu'un polype est signalé positif alors qu'il n'y a rien.
- Le **rappel** ( $=\frac{TP}{TP+FN}$ ). Le rappel permet de définir le résultat de bonnes détections par rapport au nombre total de lésions à détecter.
- La **précision** ( $=\frac{TP}{TP+FP}$ ). La précision permet de définir le résultats de bonnes détections par rapport au nombre total de lésions détectées.
- Le **F<sub>1</sub> Score** ( $=\frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$ ). Le F<sub>1</sub> Score, qui combine Précision et Rappel, est utilisé pour faire la balance entre le nombre de fausses alarmes et le nombre de lésions ratées.

| Méthode de détection | Rappel | Précision | F <sub>1</sub> Score |
|----------------------|--------|-----------|----------------------|
| Transformée de Hough | 94%    | 14%       | 24%                  |
| Log Gabor            | 42%    | 77%       | 54%                  |

TABLE 2.1 – Performances de la détection des polypes

Grâce au tableau 2.1, on constate que la transformée de Hough permet de détecter une très grande partie des polypes (rappel élevé) et que le Log Gabor permet au contraire, de bien localiser les polypes (précision élevée). Mais dans l'optique de "mieux vaut trop détecter que de rater des lésions", on va donc utiliser par la suite la transformée de Hough circulaire afin de détecter les polypes dans les images de coloscopie. On sera donc certain d'être capable de détecter le maximum de polypes possibles, moyennant l'ajout d'une étape de raffinement.

De plus, comme la transformée de Hough circulaire possède des capacités d'embarquabilité [ZIN14], elle est d'autant plus intéressante pour être utilisée pour notre problème. On va donc utiliser maintenant la chaîne de détection présentée en figure 2.7.

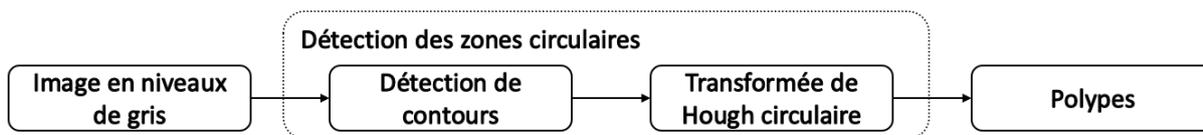


FIGURE 2.7 – Chaîne de traitement pour la détection des polypes basée sur la transformée de Hough

### Limitation de la transformée de Hough pour la détection des polypes

Comme vu précédemment dans le tableau 2.1, la transformée de Hough permet de détecter une grande quantité de polypes (rappel élevé) mais dispose d'une faible précision. Cela traduit donc un nombre élevé de fausses détections.

## 2.1. UNE MÉTHODE POUR LA DÉTECTION DES POLYPES BASÉE SUR LA FORME<sup>51</sup>

Ce nombre élevé de fausses détections est justifié par le fait que la transformée de Hough circulaire détecte justement toutes les zones circulaires sans faire de distinction en ce qui concerne le fait que la zone détectée est un polype ou non. Un regard rapide sur quelques images issues de la détection montre que la transformée de Hough détecte aussi le lumen (qui correspond à la zone noire du côlon non éclairée par le coloscope), les bulles présentes dans le côlon, la lumière spéculaire (les reflets de la source de lumière sur les parois du côlon) et quelques fois (suivant les réglages de la transformée de Hough) les bords de l'image. La figure 2.8 montre quelques exemples de résultats de détections liés à la transformée de Hough circulaire.

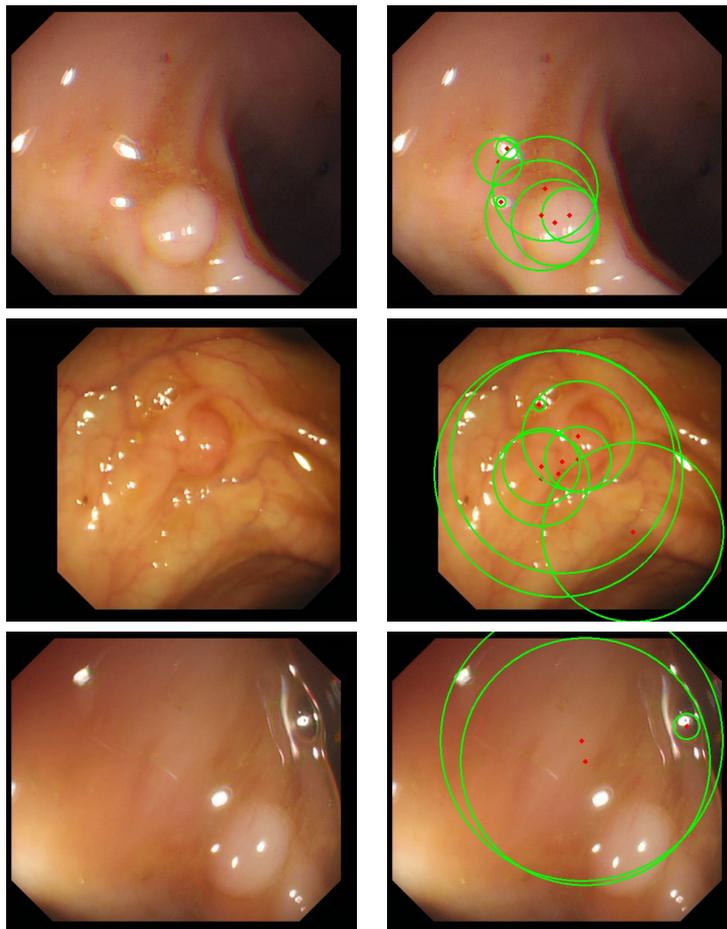


FIGURE 2.8 – Exemples de détections avec la transformée de Hough circulaire. Les images de droite montrent le résultat de la transformée de Hough sur les images originales correspondantes à gauche.

### 2.1.3 Apprentissage machine en complément

Comme on a pu le constater sur quelques exemples d'images, la transformée de Hough détecte plutôt bien les zones circulaires, mais n'est pas très sélective et détecte beaucoup

de faux positifs en ce qui concerne les polypes, c'est pour cette raison qu'on décide de proposer une nouvelle chaîne de traitement en complément.

### Une chaîne de traitement en complément de la transformée de Hough

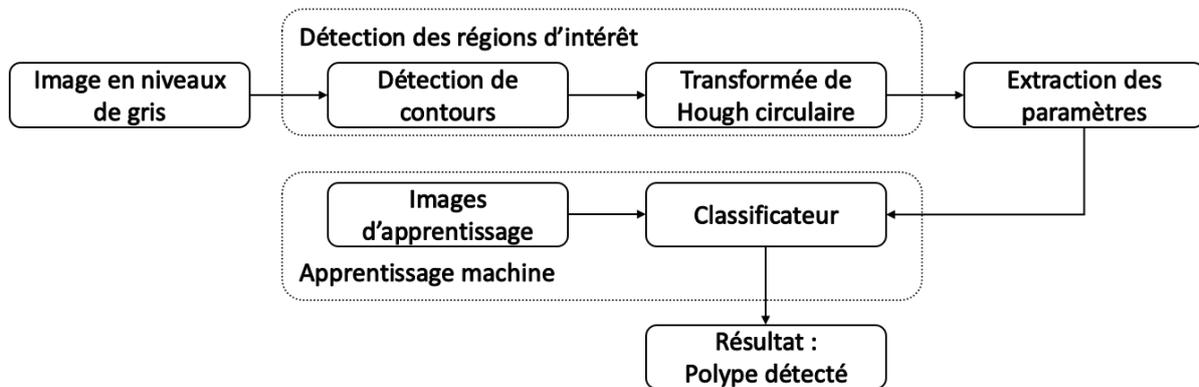


FIGURE 2.9 – Schéma de détection des polypes avec de l'apprentissage machine

Afin de palier la faible précision de la transformée de Hough, on se propose d'utiliser l'apprentissage machine. On met donc au point une nouvelle chaîne de traitement qui est décrite figure 2.9. Cette nouvelle chaîne se décompose alors en deux parties : une première qui est l'extraction de régions d'intérêt grâce à la transformée de Hough circulaire. Enfin, à partir de ces régions, on va analyser leurs structures. La structure des régions d'intérêt est obtenue en calculant les différents paramètres d'Haralick issus de la matrice de cooccurrence [HS<sup>+</sup>73]. Ces paramètres, comparés à un apprentissage effectué en amont va permettre de prédire si la zone détectée est à considérer comme polype ou non. On se propose donc dans la suite de cette partie de décrire le fonctionnement de l'apprentissage machine utilisé ainsi que l'obtention des paramètres d'Haralick, puis on s'intéressera à l'influence de cet apprentissage sur les performances de la chaîne de détection.

## Matrice de cooccurrence et paramètres d'Haralick

### Matrice de cooccurrence

La matrice de cooccurrence  $MC$  montre à quelle fréquence un pixel dont la valeur en niveau de gris  $i$  est adjacent à la valeur d'un pixel dont la valeur en niveau de gris est  $j$  (que ce soit de façon horizontale, verticale ou diagonale). La figure 2.10 montre comment obtenir la matrice de cooccurrence,  $(\Delta x, \Delta y)$  étant l'offset qui définit la relation spatiale pour laquelle cette matrice est calculée. La matrice de cooccurrence peut également être paramétrée en termes de distance  $d$  et d'angle  $\theta$  à la place de l'offset  $(\Delta x, \Delta y)$ .

La figure 2.11 montre un exemple d'une matrice de cooccurrence obtenue à partir d'une image en niveaux de gris (prise dans la base d'images Brodatz).

$$\mathcal{MC}_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1 & \text{si } I(p, q) = i \text{ et } I(p + \Delta x, q + \Delta y) = j \\ 0 & \text{sinon} \end{cases}$$

FIGURE 2.10 – Illustration du calcul de la matrice de cooccurrence

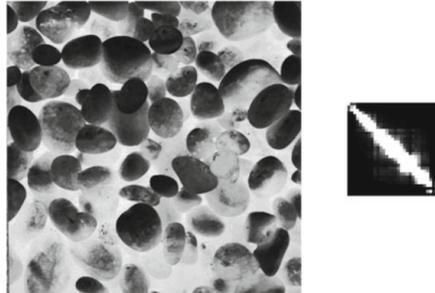


FIGURE 2.11 – Exemple d’une matrice de cooccurrence associée à son image

Le premier avantage de la matrice de cooccurrence pour l’analyse de texture pour la détection des polypes est la dimension de la matrice qui est fixée et dépend uniquement de la résolution des niveaux de gris de l’image (256, 128, 64, 32, 16 ou 8) : quelle que soit la taille de la région d’intérêt détectée, la matrice aura toujours la même taille. Cela permet de limiter l’utilisation de ressources.

Le second avantage est, qu’à partir de cette matrice, il est possible de calculer plusieurs paramètres relatifs à l’image correspondante afin d’en extraire de nouvelles informations. Les paramètres les plus utilisés avec la matrice de cooccurrence sont les paramètres d’Haralick.

### Paramètres d’Haralick

Haralick introduit plusieurs paramètres (14 au total) [HS<sup>+</sup>73] permettant d’analyser la texture de l’image. En effet, la texture d’une image est l’une des plus importantes caractéristiques qui permet d’identifier des objets ou des régions d’intérêts dans une image, quelle que soit sa provenance (ici d’une coloscope). Les paramètres d’Haralick se basent uniquement sur la matrice de cooccurrence et donc dépendent directement d’images en niveaux de gris. Les paramètres d’Haralick sont donc des paramètres qui sont calculés dans le domaine spatial et tiennent compte de la nature statistique de la texture. Ces paramètres contiennent des informations sur la texture de l’image telles que l’homogénéité, la structure linéaire, le contraste, la complexité de l’image et bien d’autres. Certains de ces paramètres sont donnés à titre illustratif dans la figure 2.12 (contraste, corrélation, entropie, second moment angulaire). Côté complexité de calcul, tout comme la complexité de calcul de la matrice de cooccurrence qui dépend des niveaux de gris de l’image, la complexité de calcul des paramètres d’Haralick dépend de la taille de la matrice, et donc plus elle est petite et plus les paramètres sont rapides à calculer.

Grâce à la combinaison de ces paramètres et de l’apprentissage sur des images d’en-

trainement, on va être capable de dire si une région d'intérêt détectée par la transformée de Hough sur des images à tester contient possiblement un polype.

$$\text{Contraste} = \frac{1}{K} \sum_{k=0}^{N-1} k^2 \sum_{|i-j|=k} \mathcal{MC}(i, j)$$

$$\text{Corrélation} = \frac{1}{K\sigma_x\sigma_y} \sum_{i,j} ij\mathcal{MC}(i, j) - \mu_x\mu_y$$

$$\text{Entropie} = -\frac{1}{K} \sum_{i,j} \mathcal{MC}(i, j) \log\left(\frac{\mathcal{MC}(i, j)}{K}\right)$$

$$\text{Second Moment Angulaire} = \frac{1}{K} \sum_{i,j} \mathcal{MC}(i, j)^2$$

Avec  $K$  le nombre d'éléments de  $\mathcal{MC}(i, j)$  et

$$\mu_x = \frac{1}{K} \sum_{i,j} i\mathcal{MC}(i, j)$$

$$\mu_y = \frac{1}{K} \sum_{i,j} j\mathcal{MC}(i, j)$$

$$\sigma_x^2 = \frac{1}{K} \sum_{i,j} (i - \mu_x)^2 \mathcal{MC}(i, j)$$

$$\sigma_y^2 = \frac{1}{K} \sum_{i,j} (j - \mu_y)^2 \mathcal{MC}(i, j)$$

FIGURE 2.12 – Exemples de paramètres d'Haralick

## Performances de l'apprentissage machine

On a pu voir précédemment que le boosting semblait être bien adapté, mais on va comparer ici ses performances par rapport à d'autres méthodes d'apprentissage. Afin de mesurer les performances de l'apprentissage machine, la base de données CVC-Colon proposée par Bernal et al. [BSV12] est utilisée, dont quelques images sont disponibles en figure 2.13.

Les images de cette base de données sont découpées de la façon suivante : des patches représentant différents endroits de l'image sont créés, et pour chaque image de la base de données, un patch contenant le polype et quatre patches n'en contenant pas sont générés. Un exemple de cette création de patches est disponible sur la figure 2.14.

La nouvelle base de données générée contient donc 1500 images (300 patches de polypes et 1200 patches d'exemples négatifs). Cette base de données est alors séparée en deux mor-



FIGURE 2.13 – Exemples d’images disponibles dans la base CVC-Colon

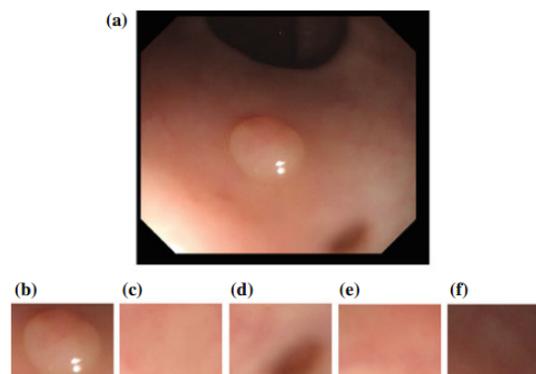


FIGURE 2.14 – Exemple de création des patches : (a) image originale, (b) exemple positif, (c-f) exemples négatifs

ceaux : le premier qui contient 1000 images (200 polypes et 800 négatifs) et qui est utilisé pour l’apprentissage machine sur les paramètres d’Haralick, et un second qui contient 500 images (100 polypes et 400 négatifs) utilisé pour mener des tests.

Afin d’être objectif sur les tests, il a été décidé de comparer différentes méthodes de classification afin de voir laquelle est la plus adaptée au problème : la méthode LVQ [Koh95] (Learning Vector Quantization, ou apprentissage pour la quantification vectorielle en français, qui peut être comprise comme un cas spécial des réseaux de neurones artificiels), la méthode d’Adaboost classique et pour finir, la méthode de Cascade Adaboost. On utilise alors la chaîne de traitement présentée en figure 2.9 afin de mesurer les performances des différentes méthodes de classification, on se propose d’utiliser les 3 métriques précédentes :

- Le rappel
- La précision
- Le  $F_1$  Score

Les résultats de ce test sont disponibles dans le tableau 2.2. Comme on peut le remarquer en regardant ces résultats, la méthode de Cascade Adaboost se démarque des deux autres en proposant un bon compromis entre rappel et précision. Même si la méthode LVQ pro-

pose un très bon rappel, sa précision reste faible par rapport à la méthode de Cascade Adaboost.

| Méthode d'apprentissage | Rappel | Précision | F <sub>1</sub> Score |
|-------------------------|--------|-----------|----------------------|
| LVQ                     | 92%    | 62%       | 74%                  |
| Adaboost                | 77%    | 72%       | 74.4%                |
| Cascade Adaboost        | 91%    | 83%       | 86.1%                |

TABLE 2.2 – Performances de la classification sur des patchs de polypes et non des images de polypes

Ainsi la méthode Cascade Adaboost reste celle qui permet d'obtenir les meilleurs résultats pour la classification des patchs de polypes. Il peut être intéressant de regarder le temps de calcul moyen mis par la méthode de détection de polype combinant la détection par la transformée de Hough suivie de la classification (dont le résultat de la détection est visible sur la figure 2.16). Cet essai, effectué sur un ordinateur disposant d'un Intel Core i7 fonctionnant à 2.8 Ghz, permet de mesurer un temps de calcul moyen de 40 millisecondes (effectué en C++). Cela traduit donc la capacité de cette méthode à détecter les polypes au rythme de 25 images par secondes, autant dire la vitesse de lecture de la plupart des vidéos existantes. De plus, cela montre que l'algorithme de détection des polypes est capable de fonctionner en temps réel durant un examen : le résultat peut être à priori affiché en parallèle au médecin sans pour autant perturber son exploration du côlon à la recherche de polypes. On obtient alors une nouvelle chaîne de traitement visible sur la figure 2.15.

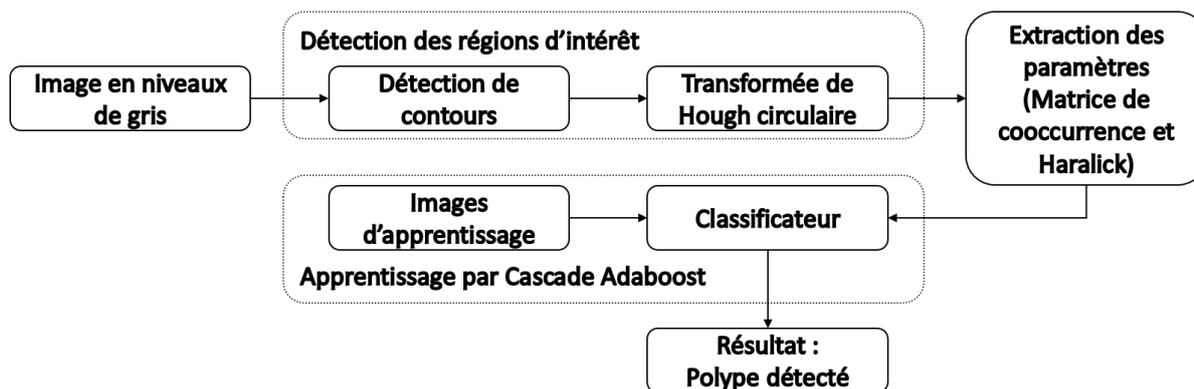


FIGURE 2.15 – Schéma de détection des polypes avec de l'apprentissage Cascade Adaboost

### 2.1.4 Conclusion et limitations de cette méthode

Cette méthode (figure 2.15) a montré qu'elle permet de détecter et localiser les polypes avec de bonnes performances et en étant capable de fonctionner en temps réel sur une

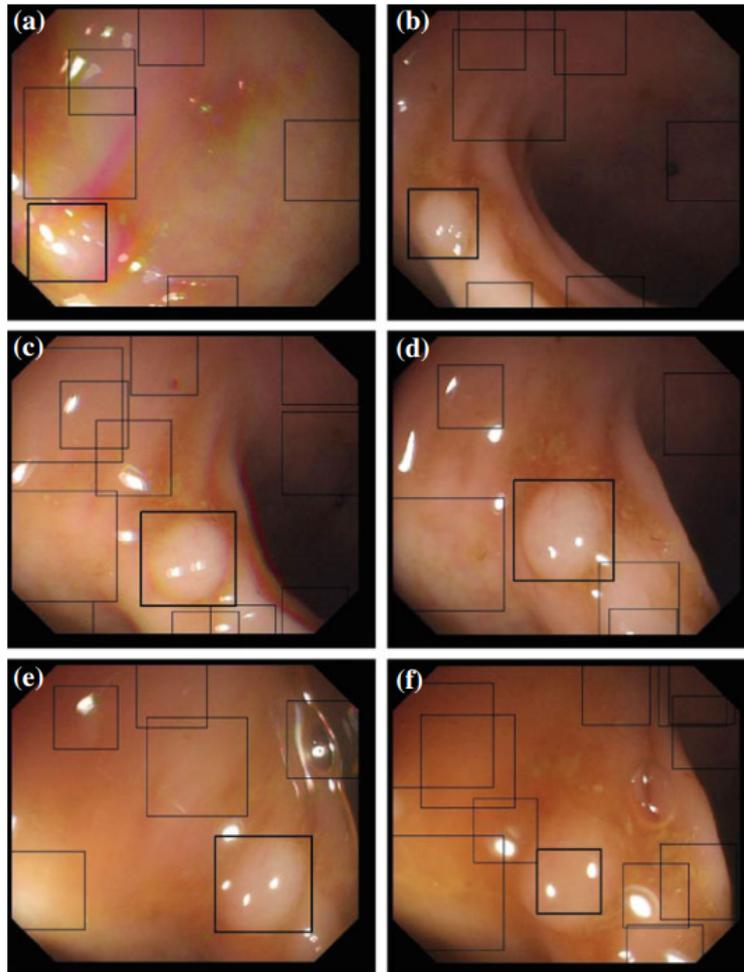


FIGURE 2.16 – Exemples de détections de polypes dans des images, les rectangles noirs sont les zones détectées par la transformée de Hough, les rectangles noirs plus épais sont les zones classées comme polype après l'étape de classification

machine de travail qui possède une configuration très performante.

Cependant, la méthode n'est pas compatible avec le changement d'échelle. Lorsque la résolution de l'image augmente, le nombre d'images que peut traiter l'algorithme en une seconde chute à 3 ou 4. Les performances sont alors très loin du temps réel.

De plus, la plupart des coloscopes disponibles dans les salles d'examen fonctionnent maintenant en haute définition d'image (HD), on peut alors s'intéresser au temps de traitement de ce type d'image sur la station de travail, et le constat est sans appel, le temps de calcul chute et on se retrouve ainsi avec un taux de traitement des images de l'ordre de 2 à 3 images par secondes. On peut donc se poser la question du résultat de ce même algorithme avec une machine moins puissante travaillant sur des images HD.

Mais plus encore, le but de cette méthode est d'aboutir à la détection lors d'examens, et donc demande de passer par la détection dans des vidéos issues de coloscopies. Or le

passage à la vidéo est beaucoup plus compliqué en termes de performances et demande une méthode plus performante sur les images avant tout. En effet, bien que la classification des patchs fonctionne, les performances de la chaîne complète sont dégradées à cause de la sensibilité de la position du patch lors de la détection par la transformée de Hough, patch qui n'est pas forcément centré sur le polype, ce qui rend donc le passage à la vidéo compliqué.

L'algorithme proposé ici, malgré ses performances de détection et caractérisation intéressantes n'est donc pas adapté à notre problème qui essaie de conjuguer performances, embarquabilité et utilisabilité clinique. C'est pour cela qu'il faut se recentrer sur autre chose sans pour autant tout oublier de ce travail. On n'oubliera surtout pas le fait que la technique d'apprentissage machine Cascade Adaboost est très prometteuse dans la classification d'éléments dans une image de coloscopie.

## 2.2 L'apprentissage actif pour la détection des polypes

Le but ici est de proposer une méthode plus adaptée que la précédente permettant de détecter les polypes selon les contraintes présentées précédemment.

Contrairement à la méthode précédente, cette méthode va alors se fonder principalement sur de l'apprentissage machine et ne sera donc plus une méthode hybride avec une préselection des zones dans l'image. En se fondant sur de l'apprentissage machine, on prend ainsi le pari que cette méthode sera compatible avec une application temps réel.

Ceci étant, l'apprentissage machine nécessite de travailler sur des paramètres extraits des images. Précédemment, on a utilisé les paramètres d'Haralick basés sur la matrice de cooccurrence, mais ils ne se sont pas révélés être un choix pertinent pour classifier les polypes, car ils traduisent principalement la texture d'une région d'intérêt, et plusieurs polypes peuvent avoir la même texture (au niveau traitement d'image) que la paroi du côlon, surtout à un stade précoce. A la place de ces paramètres, on va utiliser d'autres paramètres qui sont plus adaptés à la détection d'objets dans des images et compatible avec la technique d'apprentissage de Viola et Jones.

L'idée ici sera de robustifier le classificateur obtenu par apprentissage actif en intégrant les résultats de détection.

### 2.2.1 Implémentation matérielle

Afin de mettre techniquement en place cette méthode, on décide d'utiliser la librairie OpenCV pour effectuer l'apprentissage et le traitement d'images. En effet, OpenCV dispose d'une fonctionnalité permettant d'effectuer l'apprentissage en utilisant la méthode de Viola et Jones et dispose aussi de toutes les fonctions utiles au développement de l'outil de détection des polypes. "OpenCV (pour Open Computer Vision) est une librairie open-

source pour la vision par ordinateur et l'apprentissage machine. Elle est spécialisée dans le traitement d'images en temps réel. Elle dispose de plus de 2500 algorithmes optimisés. Ces algorithmes sont principalement utilisés pour détecter et reconnaître des visages, identifier des objets, classer des actions humaines dans des vidéos, suivre les mouvements d'une caméra, suivre des objets en mouvement, extraire des modèles 3D d'objets et bien d'autres possibilités..."

Mais c'est surtout pour ses capacités d'apprentissage machine et de détection dans des images qu'est utilisé OpenCV dans ce travail.

Pour y parvenir, l'utilisation de la fonction dédiée à l'apprentissage Cascade Adaboost est préférée en se basant sur la base de données d'images CVC-ClinicDB [BSFE<sup>+</sup>15]. Cette base de données contient 612 images de polypes d'une résolution de 384 par 288 pixels. A partir de ces images, plusieurs imageriettes sont extraites. Par image, 1 imageriette contenant le polype est extraite et 5 imageriettes ne contenant pas de polypes sont ajoutées (figure 2.17). Cela permet de construire la base d'apprentissage contenant à la fois des exemples positifs et négatifs (au total, 612 images contenant polypes et 3060 images n'en contenant pas sont disponibles). L'apprentissage est alors effectué en utilisant seulement 550 images de polypes et 3000 images d'exemples négatifs, sous la contrainte de plusieurs paramètres pour la cascade (figure 1.3) :  $f = 0.5$  (taux maximum acceptable de fausses alarmes pour étage) et  $d = 0.999$  (taux minimum de détection acceptable par étage). Cela représente alors des performances souhaitées pour la classification (une cascade à 10 étages) qui sont les suivantes :  $F_{target} = f^{10} = 0.00097$  (taux fausses alarmes global visé) et  $D_{target} = d^{10} = 0.99$  (taux de détection acceptable global visé). Cet apprentissage va générer plusieurs classificateurs suivant le type de paramètres d'apprentissage choisis (motifs binaires locaux ou paramètres pseudo-Haar et présentés par la suite). Une fois l'apprentissage effectué, il suffit d'utiliser uniquement ces classificateurs pour effectuer la détection des polypes. Ces contraintes sont nécessaires afin d'assurer une utilisabilité clinique de l'approche.

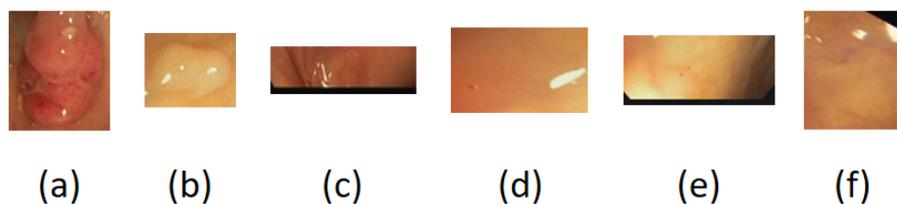


FIGURE 2.17 – Imageriettes pour la préparation de l'apprentissage, (a) (b) représentent des imageriettes avec polypes, (c) (d) (e) (f) représentent des imageriettes sans polypes

### 2.2.2 Paramètres d'apprentissage

Les paramètres à motif binaire local et les paramètres pseudo-Haar sont les deux paramètres les plus utilisés dans la littérature. Ces deux types de paramètres sont compatibles

avec la méthode d'apprentissage de Viola et Jones. On décide donc de les présenter brièvement.

### Paramètres à motif binaire local (LBP, Local Binary Pattern)

Les LBP sont des paramètres utilisés en vision par ordinateur afin d'analyser la texture d'une image ou pour effectuer la détection d'objet dans des images. Ces paramètres ont été réellement popularisés par Ojala et al. [OPH96]. Ces paramètres ont été introduits pour l'analyse de la texture d'une image en partant du principe qu'elle peut être caractérisée par un spectre de texture. Une unité de texture est représentée par huit éléments, chacun qui peut avoir une des trois valeurs possibles (0, 1, 2) obtenue par le voisinage de  $3 \times 3$  pixels. Au total, on se retrouve avec  $3^8 = 6561$  possibilités pour les unités de texture qui décrivent des motifs spatiaux à 3 niveaux dans un voisinage de  $3 \times 3$  pixels. L'occurrence de la distribution des unités de textures est calculée sur une région est appelée spectre de texture.

L'exemple de Ojala et al. [OPH96] illustre bien le fonctionnement des LBP. On peut donc expliquer plus simplement le fonctionnement des LBP en utilisant un exemple similaire. Pour cette explication, on s'appuie sur la figure 2.18. On commence donc par sélectionner un pixel (en orange sur la figure 2.18.a) et son voisinage. On effectue ensuite la différence des valeurs des pixels voisins par rapport au centre, ce qui donne la figure 2.18.b. On récupère ainsi les signes des différentes valeurs (1 si positif ou nul, 0 si négatif), on obtient donc la figure 2.18.c. Puis pour terminer, on pondère ces valeurs grâce au masque de la figure 2.18.d afin d'obtenir le résultat final figure 2.18.e. À partir de là, il suffit de sommer toutes les valeurs ( $1 + 4 + 64 + 128 = 197$ ) afin d'obtenir la valeur de l'unité de texture.

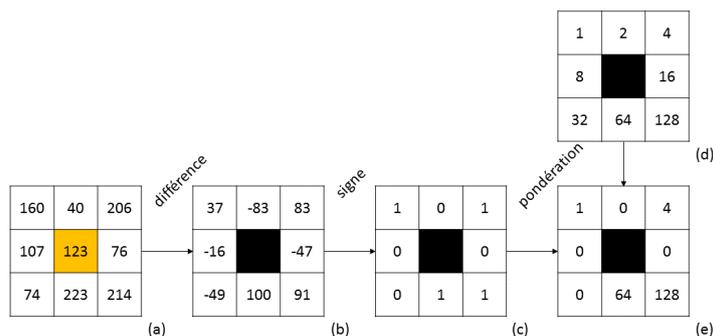


FIGURE 2.18 – Exemple d'utilisation des LBP

De plus, il existe une variante aux LBP proposée par Liao et al. [LZL<sup>+</sup>07] qui sont les MB-LBP (pour Multiscale Block LBP). Ces MB-LBP ont pour avantage d'être plus robustes que les LBP, ils permettent de traduire non seulement les microstructures des motifs de l'image mais aussi les macrostructures. Cela permet de donner une meilleure représentation de l'image que les LBP. Les MB-LBP peuvent être calculés à partir des images intégrales, d'où une rapidité d'exécution.

### Paramètres pseudo-Haar

Les paramètres pseudo-Haar sont utilisés principalement pour la détection d'objet dans des images (tout comme les LBP). Mais contrairement aux LBP, ils sont simples et rapides à calculer, ce qui permet de les utiliser pour des applications temps réel (au travers de l'utilisation de l'image intégrale). Ces paramètres ont été introduits pour la première fois par Viola et Jones [VJ01]. On considère alors des fenêtres (ou masques) de détection délimitant des zones rectangulaire adjacentes et on additionne les valeurs des pixels afin d'obtenir des sommes dont la différence va donner un paramètre. La figure 2.19 donne un exemple des premiers paramètres utilisés par Viola et Jones. Dans cet exemple, et généralement pour tous les paramètres pseudo-Haar, la somme des pixels de la zone noire est soustraite à la somme des pixels de la zone blanche.



FIGURE 2.19 – Exemples de masques proposés par Viola et Jones

En plus des masques proposés par Viola et Jones, Lienhart et Maydt ont créé des masques complémentaires [LM02] orientés de 45 degrés afin de mieux traduire les informations inclinées et donc mieux définir les contours. Quelques uns de ces masques sont disponibles figure 2.20.



FIGURE 2.20 – Exemples de masques proposés par Lienhart et Maydt

### 2.2.3 L'algorithme de base

L'algorithme utilisé pour la détection des polypes se décompose en plusieurs parties et est présenté dans le tableau 2.3. Cet algorithme permet de détecter les polypes dans des images fixes issues d'examens de coloscopie. C'est donc celui là qui sera utilisé dans un premier temps afin de mettre au point la méthode et d'affiner ses performances avant de pouvoir passer à une détection dans les vidéos.

### 2.2.4 Résultats préliminaires sur la méthode de détection

Les tests concernant la classification vont être menés sur une base de données d'images différentes de celle d'apprentissage, la base CVC-ColonDB [BSV12] déjà utilisée précédemment avec la méthode basée sur Hough. Cette base d'images contient 300 images

- 
1. Une étape de prétraitement de l'image pour préparer la détection
  2. Une étape de détection cascade qui permet de trouver les régions d'intérêt sur l'image contenant de possibles polypes en utilisant le classificateur créé précédemment (en ayant une stratégie de recherche exhaustive dans l'image)
  3. Une étape de fusion spatiale des zones détectées afin de limiter le nombre de fausses détections et ainsi réduire le nombre de régions d'intérêt affichées à l'utilisateur
  4. Une étape d'affichage permettant de retourner le résultats de la détection à l'utilisateur
- 

TABLE 2.3 – Algorithme pour la détection des polypes dans les images

de polypes. Cependant, toutes les images ne sont pas exploitables dans CVC-ColonDB, en effet, on remarque que certaines images ne sont pas nettes, par exemples, des images floues sont présentes (à cause du mouvement de l'endoscope durant la prise de vue). C'est pour cela que cette base de données a été réduite à 273 images exploitables. Des tests sont alors conduits sur différents types d'images : des images en niveaux gris (obtenues à partir de l'image couleur) et sur chacune des composantes couleurs des images (rouge, verte et bleue) afin de voir quel est le type d'image le plus intéressant pour effectuer la détection des polypes et aussi de voir si l'utilisation d'un seul canal est possible. Nous nous concentrons ici sur les LBP pour cette phase d'évaluation pour leur simplicité calculatoire et leur rapidité d'apprentissage. Les métriques suivantes sont utilisées afin de mesurer les performances de notre méthode :

- Les vrais positifs (True Positive, TP). Une détection "vraie positive" est le résultat lorsqu'un polype est détecté correctement.
- Les faux positifs (False Positive, FP). Une détection "fausse positive" est le résultat lorsqu'un polype est signalé négatif alors qu'il devrait être positif.
- Les faux négatifs (False Negative, FN). Une détection "fausse négative" est le résultat lorsqu'un polype est signalé positif alors qu'il n'y a rien.
- Le rappel ( $= \frac{TP}{TP+FN}$ ). Le rappel permet de définir le résultat de bonnes détections par rapport au nombre total de lésions à détecter.
- La précision ( $= \frac{TP}{TP+FP}$ ). La précision permet de définir le résultats de bonnes détections par rapport au nombre total de lésions détectées.
- Le  $F_1$  Score ( $= \frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$ ). Le  $F_1$  Score, qui combine Précision et Rappel, est utilisé pour faire la balance entre le nombre de fausses alarmes et le nombre de lésions ratées.
- Le **temps de détection moyen pour une image**. Il s'agit du temps moyen (moyenne prise sur l'ensemble des 273 images) mis par la méthode pour effectuer la détection des polypes dans une image.

Les résultats de ce test sont disponibles dans le tableau 2.4. Ces résultats sont obtenus en combinant la méthode du centroïde (qui consiste à classer comme positif une région d'intérêt si son centre de gravité est situé sur la vérité terrain) avec la méthode de recou-

| Classificateur LBP                | Niveaux de gris | Canal rouge | Canal vert | Canal bleu |
|-----------------------------------|-----------------|-------------|------------|------------|
| Rappel                            | 56.78%          | 87.18%      | 88.24%     | 93.04%     |
| Précision                         | 56.99%          | 21.24%      | 21.16%     | 19.23%     |
| F <sub>1</sub> Score              | 56.88%          | 34.15%      | 34.13%     | 31.87%     |
| Temps de détection pour une image | 221ms           | 92ms        | 66ms       | 51ms       |

TABLE 2.4 – Résultats de la détection sur images fixes

vrement (qui consiste à classer comme positif une région d'intérêt si sa surface recouvre un certain pourcentage de la surface du polype). On constate alors que le canal bleu permet d'obtenir le meilleur taux de détection de polypes (93.04%), résultat qui confirme les résultats de J. Bernal [BSV13] montrant les avantages du canal bleu dans la détection des polypes. En effet, J. Bernal montre que l'utilisation du canal bleu de l'image couleur permet de réduire l'information liée aux vaisseaux sanguins et permet donc d'améliorer les contours du polype, et donc en facilite sa détection. Ce même canal bleu permet d'obtenir le temps de calcul pour une image qui est le plus faible est des 4 configurations testées. Cependant, on note qu'avec ce canal, la précision est assez faible (19.23%) qui traduit donc un nombre de fausses détections assez élevé. La diminution du temps de calcul s'explique par le fait que la Cascade Adaboost élimine de fausses détections dans ses premiers étages (sur un total de 10). Il y a donc moins de régions d'intérêt qui parcourent la totalité de la Cascade, d'où un temps de calcul amélioré.

On choisit donc d'utiliser le canal bleu par la suite pour la mise au point de la méthode car il permet d'obtenir un temps de calcul assez faible tout en conservant un taux de détection assez élevé. Il faut néanmoins trouver un moyen permettant de réduire le nombre de fausses détections.

### 2.2.5 L'apprentissage actif : une solution pour réduire les fausses détections

#### Principe de l'apprentissage actif

Maintenant que l'on dispose de ce nouvel outil et des performances de ce dernier, on peut effectuer le constat suivant : il reste encore un nombre élevé de fausses détections malgré l'utilisation de la technique "Cascade Adaboost". Il faut donc trouver une nouvelle solution qui va permettre de répondre à ce problème et réduire ce nombre de fausses détections sans pour autant dégrader le Rappel. C'est pour cela qu'on se propose d'améliorer l'apprentissage en utilisant une phase d'apprentissage actif en amont de la détection.

De façon générale, l'apprentissage actif est un cas spécial de l'apprentissage machine dans lequel le processus est capable d'interagir avec l'utilisateur pour atteindre le but

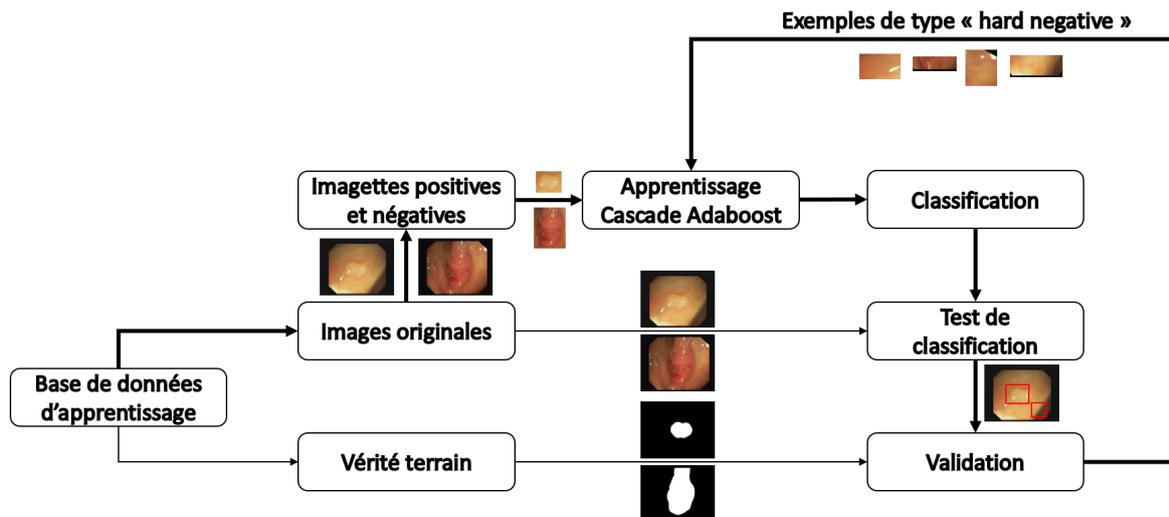


FIGURE 2.21 – Schéma de principe de l'apprentissage actif

désiré. Dans le cas de la détection des polypes, qui est beaucoup plus spécifique, l'apprentissage actif est utilisé pour améliorer la classification, qui après une première phase d'apprentissage n'obtient pas de performances suffisamment satisfaisantes. Le processus suivant (illustré avec la figure 2.21 et présenté dans le tableau 2.5) a donc été mis en place pour répondre à cette problématique [AHR16].

- 
1. A partir de la base de données d'apprentissage (CVC-ClinicDB [BSFE<sup>+</sup>15]), un premier classificateur est construit sur la base de patches
  2. Cette classification est ensuite testée sur la base de données d'apprentissage pour intégrer la notion de robustesse
  3. Les résultats de ce test sont comparés avec la vérité terrain disponible avec la base de données
  4. Toutes les détections de patches classées au final comme "faux positifs" deviennent alors de nouveaux exemples d'apprentissage de type "hard negative"
  5. Une nouvelle étape d'apprentissage est effectuée avec les images de la base d'apprentissage plus les images de type "hard negative"
- 

TABLE 2.5 – Principe de l'apprentissage actif

Ainsi, avec cette technique, le nombre de fausses détections devrait être réduit. Et afin d'améliorer encore plus la classification, il est possible de répéter les étapes 2 à 5 du tableau 2.5 plusieurs fois.

### Effets de l'apprentissage actif sur les performances

Ici, le processus d'apprentissage actif a été effectué trois fois en se focalisant sur les LBP et les résultats sont disponibles dans le tableau 2.6. Il est alors possible de constater que les résultats de la détection s'améliorent avec les répétitions du processus d'apprentissage actif : le  $F_1$  Score passe ainsi de 31.87% à 47.55%. De plus, le temps de calcul par image chute de 51 millisecondes à 39 millisecondes, qui est inférieur aux 40 millisecondes requises afin de traiter 25 images par secondes. Il y a donc un intérêt non négligeable à utiliser l'apprentissage actif ici.

Le nombre de répétitions du processus d'apprentissage actif (ici trois répétitions) n'a pas été choisi au hasard : quand on observe la figure 2.22, le  $F_1$  Score tend vers une asymptote. Cela semble donc indiquer que l'apprentissage actif a une limite. Puis si on se rapproche de cette limite, on risque de sur-apprendre et d'être trop spécifique : on pourrait craindre que les performances se détériorent. De plus, comme les temps d'apprentissage augmentent avec le nombre de répétitions, il est préférable de s'arrêter à ce nombre de répétitions pour se concentrer sur le passage à la vidéo à la vue des temps de calcul (39 millisecondes d'après le tableau 2.6) et de combler les performances avec de nouvelles solutions.

| Classifieur LBP                   | Sans A.A. | Avec A.A. 1 fois | Avec A.A. 2 fois | Avec A.A. 3 fois |
|-----------------------------------|-----------|------------------|------------------|------------------|
| Rappel                            | 93.04%    | 93.77%           | 88.28%           | 86.21%           |
| Précision                         | 19.23%    | 23.66%           | 30.70%           | 32.83%           |
| $F_1$ Score                       | 31.87%    | 37.78%           | 45.55%           | 47.55%           |
| Temps de détection pour une image | 51ms      | 44ms             | 40ms             | 39ms             |

TABLE 2.6 – Résultats de la détection sur images fixes avec l'apprentissage actif (A.A.)

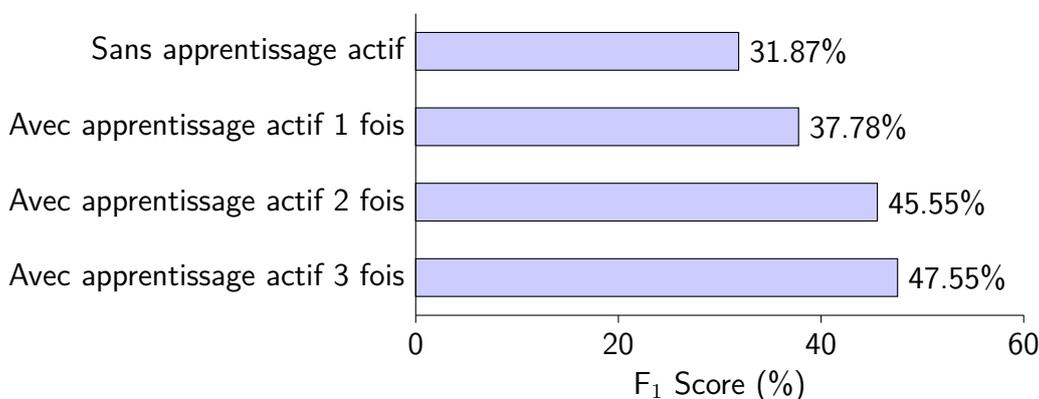


FIGURE 2.22 – Évolution du  $F_1$  Score avec les répétitions d'apprentissage actif

### Conclusion sur l'apprentissage actif

On a ainsi pu voir ici que l'apprentissage machine pouvait donner de bons résultats. Cependant, comme la quantité de données disponible pour effectuer l'apprentissage reste limitée, il a fallu mettre au point une technique permettant de l'améliorer. Cette amélioration a été obtenue en mettant en place un processus d'apprentissage actif. Cet apprentissage actif se base sur les résultats de classification après l'obtention du premier classificateur. A partir de là, il est capable d'extraire toutes les zones détectées erronées (c'est-à-dire tous les faux positifs) afin de créer de nouveaux exemples négatifs pour l'apprentissage. Puis on génère un nouveau classificateur comprenant ces nouveaux exemples. Il est alors possible de répéter ce processus plusieurs fois afin d'en améliorer les performances. On a ainsi pu voir que les performances s'amélioraient (l'apprentissage actif permet bien de réduire les fausses détections) et que le temps de détection chutait sous la barre des 40 millisecondes.



FIGURE 2.23 – Résultats de la méthode avec apprentissage actif

Comme les résultats obtenus sont corrects et que le temps de traitement moyen par image est en dessous de 40 millisecondes, on peut s'intéresser maintenant aux performances de la détection dans des vidéos issues de coloscopie. Un exemple visuel des résultats obtenus avec cette méthode sont disponibles figure 2.23.

## 2.3 Conclusion sur la détection dans les images

On a donc pu voir dans cette partie une première méthode, permettant de détecter les polypes dans les images issues de coloscopies, basée sur une approche géométrique (la détection des formes circulaires) et associée à de l'apprentissage machine pour suppléer la détection des formes circulaires et améliorer la détection des polypes. En effet, la transformée de Hough avait l'avantage de détecter une majorité de polypes, mais entraînait un nombre élevé de fausses détections (cela s'est vu avec la précision faible). L'apprentissage machine (ici la méthode de Viola et Jones) permet théoriquement d'améliorer ces résultats. Cependant, cette méthode montre ses limites en termes de temps de calcul : la matrice de cooccurrence dépend de la résolution de l'image, et donc, plus la résolution augmente, plus le temps de calcul augmentera. On s'éloigne alors des performances proches du temps réel. Et comme les coloscopes actuels ont une image en haute définition, cette méthode hybride n'est pas adaptée.

Cependant, cette première méthode a montré quelque chose d'intéressant : les performances de détection et classification en utilisant la méthode d'apprentissage de Viola et Jones (ou encore appelée la méthode de Cascade Adaboost). On peut alors s'intéresser à mettre au point un outil uniquement basé sur l'apprentissage machine en utilisant la méthode de Viola et Jones : cela a donné naissance à la seconde méthode.

La seconde méthode (utilisant d'autres paramètres que ceux d'Haralick) va donc permettre une meilleure détection des polypes. Plusieurs tests ont été menés dans un premier temps pour identifier le canal de l'image qui apportait le meilleur taux de détection, qui s'est révélé être le canal bleu, en accord avec des résultats de la littérature. Cependant, le nombre de fausses détections était encore élevé, c'est pour cela qu'on a mis au point une méthode d'apprentissage actif, qui permet de réinjecter des exemples négatifs dans l'apprentissage afin d'en améliorer les performances. On a ainsi pu voir que cette méthode portait ses fruits en proposant des résultats intéressants en termes de taux de détection, mais qu'en plus, elle permettait de réduire le temps de calcul à moins de 40 millisecondes. Cependant, on a aussi vu que l'apprentissage actif montrait ses limites après 3 étapes pour les LBP.

Si on se réfère aux contraintes fixées, la seconde méthode permet de répondre à 3 des 4 points requis dans notre cahier des charges. En effet, on est capable de traiter les images en temps réel (39 millisecondes par image, soit moins des 40 millisecondes requises), la méthode est embarquable car toutes les méthodes utilisées pour réaliser l'outil le sont et les performances de détection de notre outil sont acceptables ( $F_1$  Score de 47.55%), au regard d'un passage à la vidéo.

Pour confronter le dernier point, c'est-à-dire la compatibilité avec une utilisation clinique de notre outil, on va se rapprocher de vrais cas d'examens, et donc s'intéresser à la détection des polypes dans les vidéos issues de coloscopies.



## Chapitre 3

# La détection des polypes dans les vidéos issues de coloscopies

Comme on a pu le voir précédemment, le temps de calcul pour traiter une image fixe est de moins de 40 millisecondes, ce qui traduit une capacité à traiter toutes les images à la vitesse minimale de 25 images par seconde. On peut donc s'intéresser à la capacité de notre outil à détecter des polypes dans des vidéos issues de coloscopies, et surtout s'intéresser à la compatibilité avec une utilisation clinique de notre outil.

En faisant des essais sur des vidéos, il est alors possible de se rapprocher des cas réels de coloscopie et de prouver l'intérêt de la méthode de détection des polypes pour des examens en conditions hospitalières.

Mais il ne faut pas uniquement se concentrer sur le dernier point de la compatibilité avec une utilisation clinique, on garde toujours en mémoire toutes les contraintes de notre outil, que l'on rappelle ici :

- être capable de détecter les polypes en temps réel (c'est-à-dire que le temps pour traiter une image issue de coloscopie doit être inférieur à 40 millisecondes)
- être compatible avec du matériel dont les ressources de calculs sont standards et limitées
- obtenir des performances élevées en termes de sensibilité et de spécificité
- être compatible avec l'utilisation en routine clinique (et donc ne pas modifier la pratique clinique du médecin) et alerter le médecin sur la présence de polypes durant l'examen.

Dans cette partie, on va donc se consacrer aux performances dans les vidéos, et pour cela, on va introduire dans un premier temps de nouvelles métriques permettant de mieux traduire ces performances. On introduit également une nouvelle base de données contenant des vidéos de coloscopies. Par la suite, on présentera les résultats de cette méthode sur cette base de données puis les différentes optimisations mises en place et possibles afin d'améliorer les performances de l'outil.

## 3.1 La détection dans des vidéos

### 3.1.1 Introduction de nouvelles métriques pour la mesure de performances dans les vidéos

Tout comme avec les images fixes, il faut des métriques pour mesurer les performances de notre outil. Précédemment, afin de classer un résultat de détection sur une image, on s'est basé sur la méthode dite du centroïde, c'est-à-dire que pour classer un résultat, on regardait sur la vérité terrain si le centre de la région d'intérêt détectée se situait bien sur la zone contenant le polype. Mais cette méthode est stricte et surtout, ne rend pas compte de l'aspect visuel de la détection, qui ne traduit donc pas le fait d'attirer l'attention du médecin sur la zone de l'image contenant le polype à détecter dans une vidéo. On se propose alors, en plus d'utiliser une partie des métriques présentées précédemment (Précision, Rappel et  $F_1$  Score), d'introduire de nouvelles métriques utilisables cliniquement [ABSM<sup>+</sup>17a], qui permettent de mieux traduire les performances de la méthode d'un point de vue clinique, et donc de mieux interpréter les résultats de la méthode utilisée (il s'agit ici d'être plus clair pour les médecins). Ces nouvelles métriques sont alors les suivantes :

- Le **taux de détection des polypes (Polyp Detection Rate, PDR)**. Il permet de montrer si la méthode a été capable de détecter le polype au moins une fois dans la séquence.
- Le **temps de calcul moyen par image (Mean Processing Time, MPT)**. Il correspond au temps moyen mis pour traiter une image. Il est associé à la contrainte temps réel pour l'utilisation clinique. Si on considère que les vidéos sont enregistrées à 25 images par secondes, alors la contrainte temps réel est de 40 millisecondes maximum.
- Le **nombre moyen de faux positifs par image (Mean Number of False Positive, MNFP)**. Il s'agit avec cette mesure de savoir si la méthode utilisée pour détecter les polypes ne perturbe pas l'attention du médecin en affichant de fausses détections.
- Le **temps de réaction (Reaction Time, RT)**. Cela représente le délai entre la première apparition d'un polype dans la vidéo et la première bonne détection de ce polype par la méthode utilisée, mesure effectuée à la fois en nombre d'images et en secondes (si on considère la vidéo défilant à 25 images par secondes). La figure 3.1 illustre comment est calculé le temps de réaction.

Ce sont donc ces métriques que l'on va utiliser par la suite pour mesurer les performances de notre outil et comparer les résultats des différentes optimisations par rapport à la méthode de base.

Afin de pouvoir calculer les différentes métriques (et en particulier le Rappel et la Précision), il faut pouvoir classer correctement les différentes détections renvoyées par le système, c'est-à-dire si une région d'intérêt détectée se classe comme vraie positive, fausse positive, fausse négative ou vraie négative. Pour cela, nous décidons de mettre au point un nouveau processus permettant de classer au mieux les zones détectées. Ce processus

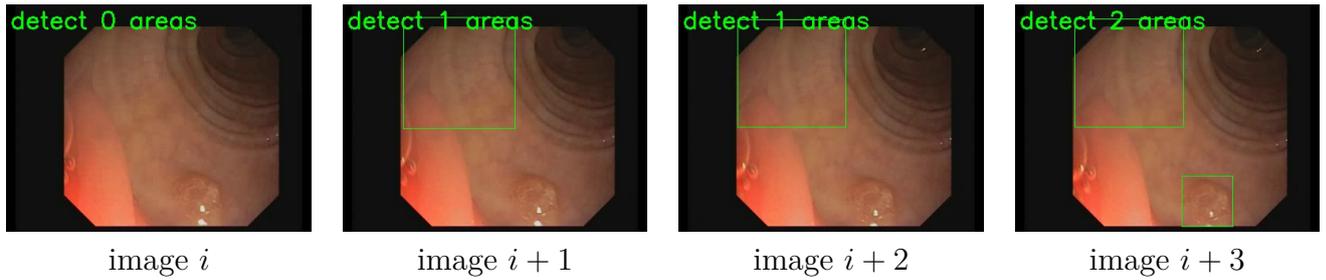


FIGURE 3.1 – Illustration du temps de réaction. Le polype apparaît pour la première fois à l'image  $i$  et est correctement détecté à l'image  $i + 3$ , il s'est donc écoulé un total de 4 images pour avoir une détection correcte après la première apparition du polype, donc un temps de réaction de 160 millisecondes ( $4 \times 40ms = 160ms$ )

est présenté dans la figure 3.2. En effet, précédemment, pour classer une détection, on se basait sur la méthode dite du centroïde : on comparait la position du centre de la région d'intérêt détectée par rapport à la vérité terrain. Mais cette méthode était trop stricte et ne traduisait pas l'aspect suivant : attirer l'oeil du médecin sur le polype. Avec le processus présenté ici, on tient compte de cela.

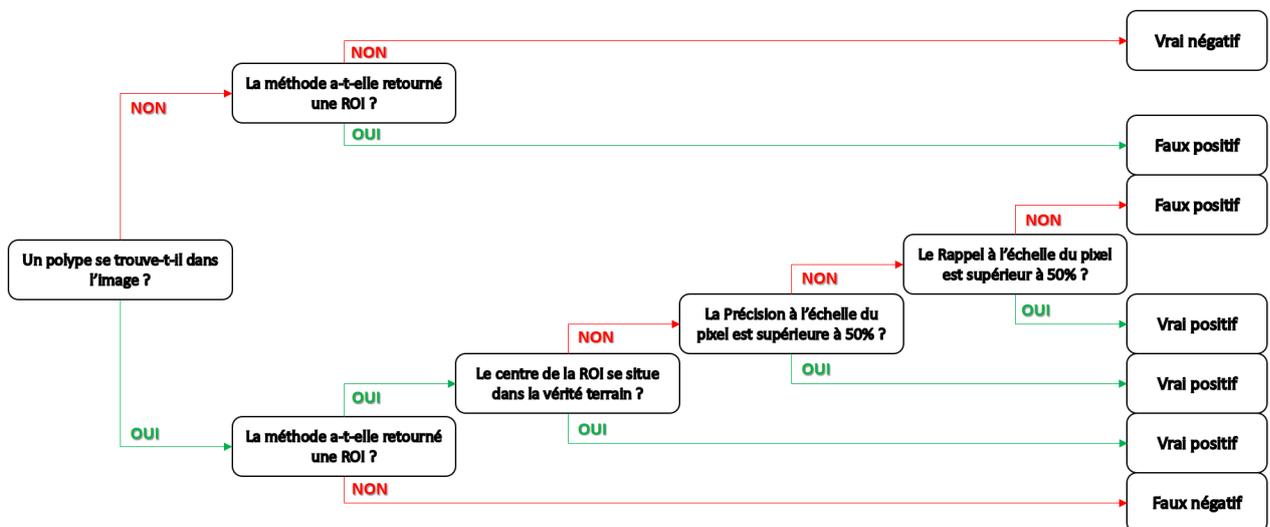


FIGURE 3.2 – Processus de classification d'une région d'intérêt (ROI)

Ce processus fonctionne à deux échelles : d'abord à l'échelle de l'image, puis à l'échelle du pixel. En s'appuyant sur la vérité terrain, on sait si un polype est présent ou non sur l'image traitée. On va ensuite enchaîner les différentes étapes de notre arbre de décisions afin de savoir si la région d'intérêt détectée par notre outil est à considérer comme vrai positif, vrai négatif, faux positif ou faux négatif. C'est donc en se basant sur les résultats obtenus, à travers ce processus, qu'on va pouvoir calculer les différentes métriques qui vont refléter les performances de notre méthode de détection des polypes.

De plus, l'intérêt de proposer un nouveau processus afin de mieux classer les régions d'intérêt détectées est d'accorder plus d'importance à l'aspect visuel de la détection et de traduire indirectement la capacité de la méthode à attirer l'oeil du médecin sur le polype sans nécessairement encadrer complètement le polype.

### Base de données vidéo

Afin de pouvoir tester notre méthode, on se doit de disposer d'une base de données composée de vidéos issues de coloscopies et non plus seulement d'images fixes. C'est pour cela qu'on décide d'utiliser celle présentée dans [ABSM<sup>+</sup>17b]. Cette base de données dispose de 18 vidéos uniques extraites de coloscopies autour de polypes et de la vérité terrain associée pour chaque image des vidéos, il sera alors possible de se rendre compte des qualités de détection de notre méthode. Ces vidéos possèdent une résolution de 768 pixels par 576 pixels et ont été obtenues grâce à la combinaison d'un endoscope Olympus QF190 et d'un système de numérisation d'image vidéo Exera III. La figure 3.3 représente le profil de la vidéo numéro 18, c'est-à-dire la présence ou non d'un polype à l'écran (0 si le polype est absent, 1 s'il est présent). Plus loin en annexe, les figures A.1.a, A.2.a, A.3.a, A.4.a, A.5.a, A.6.a, A.7.a, A.8.a, A.9.a, A.10.a, A.11.a, A.12.a, A.13.a, A.14.a, A.15.a, A.16.a, A.17.a et A.18.a représentent les profils des 18 vidéos qui permettent de s'apercevoir qu'elles représentent des cas bien réels et variés. Au total, on dispose donc de 11 954 images afin de mener les tests.

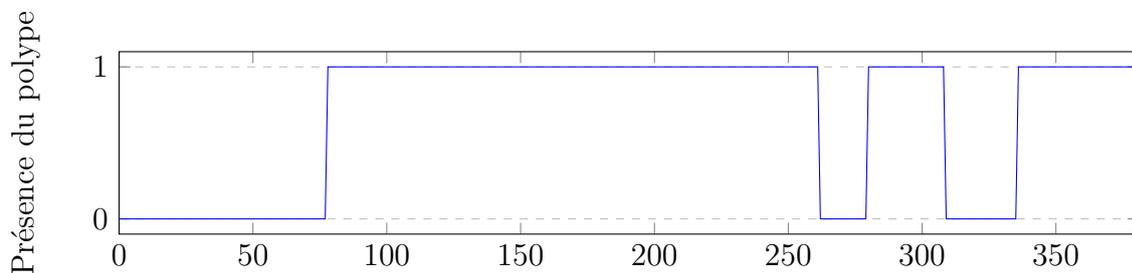


FIGURE 3.3 – Profil de la vidéo 18 (381 images)

Sur ces données, on a décidé d'aller plus loin en classant qualitativement ces 18 vidéos selon qu'il est plus ou moins difficile de détecter le polype. Les critères suivants ont été utilisés :

- La taille du polype, plus il est gros, plus il sera facile de le détecter (et vice versa)
- La netteté de l'image
- La présence de matière non désirée sur l'image (résidus, bulles)

Cela permet alors de classer ces vidéos en leur accordant une note allant de 0 à 3, 0 indiquant que le polype est facile à détecter, 3 indiquant qu'il est difficile à détecter dans

la vidéo. Cette classification est disponible dans le tableau 3.1.

On s’aperçoit donc que certains polypes sont plus difficiles que d’autres à détecter dans les vidéos disponibles dans cette base de données (environ 28%).

|                      |           |           |           |           |           |           |           |           |           |
|----------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Numéro de vidéo      | <b>1</b>  | <b>2</b>  | <b>3</b>  | <b>4</b>  | <b>5</b>  | <b>6</b>  | <b>7</b>  | <b>8</b>  | <b>9</b>  |
| Niveau de difficulté | 1         | 1         | 0         | 0         | 1         | 0         | 2         | 2         | 0         |
| Numéro de vidéo      | <b>10</b> | <b>11</b> | <b>12</b> | <b>13</b> | <b>14</b> | <b>15</b> | <b>16</b> | <b>17</b> | <b>18</b> |
| Niveau de difficulté | 1         | 1         | 1         | 2         | 2         | 2         | 1         | 1         | 0         |

TABLE 3.1 – Niveaux de difficulté de détection des polypes dans les vidéos

### Protocole de mesure

Afin de mesurer de manière adaptée les performances de notre méthode, on a mis en place un protocole pour avoir des mesures comparables entre elles. Tout d’abord, côté mesure des performances de détection, on utilise le processus de classification d’une région d’intérêt détectée (figure 3.2) ainsi que toutes les métriques classiques présentées jusque là en se basant sur les résultats du dit processus. En ce qui concerne la mesure des temps de calcul, on prend la moyenne des temps de calcul sur les 18 vidéos après 10 exécutions. Cela permet de rendre moins sensible les mesures aux variations de charge de travail que peut subir l’ordinateur sur lequel on effectue nos tests. Concernant la machine utilisée, il s’agit d’une station de calculs de marque Dell ayant les caractéristiques suivantes : un processeur Intel Xeon E5-1603 v3 (fonctionnant à une fréquence de 2.80 Giga-Hertz), une mémoire vive de 32 Giga-Octets et tournant sous Ubuntu 14.04 LTS (64 bits). L’outil a été développé en utilisant le langage Python, interprété dans sa version 2.7.6 et utilisant la version 3.1.0 d’OpenCV.

### 3.1.2 Résultats préliminaires de détection dans les vidéos

Avec ces métriques et cette base de données, on est donc capable de mener des tests préliminaires afin de regarder les performances de notre méthode. La méthode développée est décrite dans le tableau 3.2. Cette méthode permet donc de détecter les polypes dans les vidéos de coloscopies en utilisant les deux types de caractéristiques utilisées pour l’apprentissage : les LBP et les caractéristiques pseudo-Haar.

Quand on parle par la suite des mesures de performances de notre outil (en termes de détection et de temps de calcul), il s’agit de l’évaluation de la totalité des différentes étapes présentées dans le tableau 3.2 pour aboutir au résultat final, c’est-à-dire la détection (ou non) d’un polype.

Maintenant que l’on dispose d’une base de données, de métriques adaptées et de la méthode à tester, nous proposons des résultats préliminaires sur cette dernière.

- 
1. Une étape de prétraitement de l'image pour préparer la détection (extraction de canal)
  2. Une étape de détection cascade à différentes échelles qui permet de trouver les régions d'intérêt sur l'image contenant de possibles polypes
  3. Une étape d'affichage permettant de retourner le résultats de la détection à l'utilisateur
- 

TABLE 3.2 – Algorithme de détection des polypes dans les vidéos

## Résultats

Comme vu juste avant, on va effectuer une évaluation préliminaire des performances de la méthode. Pour y parvenir, on va donc utiliser la base de données présentée ainsi que les différentes métriques disponibles. Côté méthode de détection, on va utiliser les deux types de classifications qui sont à notre disposition : la classification utilisant les motifs locaux binaires (LBP) et la classification utilisant les caractéristiques pseudo-Haar effectuée sur la base de données CVC-Clinic (comme dans la chapitre précédent). Pour ces deux types de classification et pour le test, on insiste sur le fait qu'aucune optimisation n'est utilisée. Autrement dit, on n'utilise pas l'apprentissage actif ici car on cherche juste à s'intéresser aux performances brutes de notre méthode. Les résultats sont alors disponibles dans le tableau 3.3.

| Méthode | PDR  | MPT   | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT         |
|---------|------|-------|------|-----------|--------|----------------------|------------|
| LBP     | 100% | 140ms | 9.9  | 4.48%     | 55.33% | 8.29%                | 4.8 (0.2s) |
| Haar    | 100% | 26ms  | 5.8  | 9.20%     | 70.65% | 16.27%               | 6.8 (0.3s) |

TABLE 3.3 – Résultats préliminaires de la détection dans les vidéos sans optimisations

Suivant les résultats du tableau 3.3, on s'aperçoit que quelque soit le type de classification utilisé, on est capable d'attirer l'attention du médecin sur chacun des polypes présents dans les vidéos (ce qui correspond au PDR de 100%). Côté temps de calcul, on constate que la classification avec les caractéristiques pseudo-Haar est capable de détecter les polypes en temps réel (MPT de 26 millisecondes), mais aussi propose de meilleures performances de détection que la classification utilisant les motifs locaux binaires. Cependant, à l'inverse, on peut noter que la classification utilisant les motifs locaux binaires est plus rapide pour trouver le polype après sa première apparition dans la vidéo (cela se voit en comparant les temps de réaction entre les deux types de classification).

De plus, quand on compare les temps de calcul pour les LBP par rapport à ceux du chapitre précédent, on s'aperçoit qu'ils sont différents. Cela s'explique par le fait qu'avant, on effectuait la détection à un seul facteur d'échelle, ici, on l'effectue à 3 facteurs d'échelle différents. On a décidé de le faire à 3 échelles différents car on s'est aperçu qu'avec la classification utilisant les paramètres pseudo-Haar, les temps de calcul étaient très faibles,

et qu'en faisant la détection à 3 facteurs d'échelles différents, on rend notre méthode plus robuste pour les différentes tailles de polypes.

On voit donc ici qu'on obtient des résultats intéressants ( $F_1$  Score de 8.29% pour la classification utilisant les LBP et de 16.27% pour celle utilisant les paramètres pseudo-Haar) mais qui peuvent être améliorés par plusieurs optimisations. On va donc s'intéresser par la suite à toutes ces améliorations possibles qui vont permettre d'obtenir de meilleures performances pour la détection des polypes colorectaux dans les vidéos de coloscopies.

## 3.2 Amélioration des performances dans les vidéos

Dans cette partie, on va s'intéresser à toutes les améliorations qu'il est possible de mettre en place afin d'améliorer les performances de la méthode. Parmi ces améliorations, il y a bien évidemment l'apprentissage actif, qui permet d'obtenir de meilleures performances comme on a pu le voir avec les images fixes. Mais on va aussi introduire d'autres améliorations qui s'appliquent bien à la vidéo (et non plus aux images fixes) comme la cohérence temporelle, la fusion spatiale des zones détectées et des étapes de prétraitements.

### 3.2.1 Fusion spatiales des zones détectées

On va ici s'intéresser à la première optimisation mise en place afin d'améliorer les résultats de la détection. A la vue du grand nombre de fausses détections de la méthode, cette première optimisation va consister à fusionner spatialement les zones les plus proches afin de réduire le nombre de régions d'intérêt renvoyé par la méthode. On a donc un nouvel algorithme mis en place, disponible dans le tableau 3.4. Ce critère de fusion se base sur deux éléments : le nombre de régions d'intérêt proche et le recouvrement de ces zones entre elles, dans notre cas, on a décidé de fixer ces paramètres à 2 (car on effectue 3 détections à différents facteurs d'échelles) et avec un taux de recouvrement de 70% entre ces zones (car si on prenait en dessous de cette valeur, on ne serait pas assez sélectif, et au dessus, on deviendrait trop sélectif). On teste donc les avantages de cette fusion sur l'ensemble des 18 vidéos. Les résultats sont disponibles dans les tableaux 3.5 et 3.6.

D'après les résultats obtenus (et présentés dans les tableaux 3.5 et 3.6), la fusion spatiale des zones détectées permet d'améliorer nettement les résultats de notre outil sans en dégrader certains aspects. En effet, le temps de calcul moyen pour traiter une image ne change presque pas, donc cet aspect qui nous est essentiel n'est pas affecté. Côté détection, on est toujours capable d'alerter le médecin sur la présence des polypes (PDR de 100%) mais en n'envoyant beaucoup moins de fausses informations (au regard du MNFP), donc l'information renvoyée au médecin est beaucoup plus pertinente et claire. Au niveau performances de détection, le  $F_1$  Score augmente nettement (on passe du simple au double), certes on détecte moins de polype mais on est plus précis sur la détection (au regard des valeurs du rappel et de la précision), ce qui est un aspect important pour le médecin. La seule variable qui se dégrade est le temps de réaction qui augmente (mais reste inférieur à 1 seconde), malgré l'avantage de réduire les fausses détections, la fusion spatiale va aussi

1. Une étape de prétraitement de l'image pour préparer la détection (extraction de canal)
2. Une étape de détection cascade à différentes échelles qui permet de trouver les régions d'intérêt sur l'image contenant de possibles polypes
3. Une étape de fusion spatiale des zones détectées afin de limiter le nombre de fausses détections et ainsi réduire le nombre de régions d'intérêt affichées à l'utilisateur
4. Une étape d'affichage permettant de retourner le résultats de la détection à l'utilisateur

TABLE 3.4 – Algorithme de détection des polypes dans les vidéos avec fusion spatiale

| Fusion | PDR  | MPT   | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT         |
|--------|------|-------|------|-----------|--------|----------------------|------------|
| Sans   | 100% | 140ms | 9.9  | 4.48%     | 55.33% | 8.29%                | 4.8 (0.2s) |
| Avec   | 100% | 140ms | 3.2  | 12.41%    | 54.58% | 20.22%               | 7.2 (0.3s) |

TABLE 3.5 – Comparaison des performances de la méthode avec fusion spatiale pour la classification utilisant les motifs locaux binaires (LBP)

| Fusion | PDR  | MPT  | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT          |
|--------|------|------|------|-----------|--------|----------------------|-------------|
| Sans   | 100% | 26ms | 5.8  | 9.20%     | 70.65% | 16.27%               | 6.8 (0.3s)  |
| Avec   | 100% | 26ms | 1.4  | 23.29%    | 46.82% | 31.10%               | 17.5 (0.7s) |

TABLE 3.6 – Comparaison des performances de la méthode avec fusion spatiale pour la classification utilisant les paramètres pseudo-Haar

faire disparaître des bonnes détections, ce qui va alors augmenter le temps de réaction sur certaines vidéos.

Par la suite, on se basera donc sur cet algorithme (présenté tableau 3.4) qui sert de base dans notre méthode. De plus, c'est sur celui là qu'on appliquera les prochaines améliorations.

### 3.2.2 Apprentissage actif

Ici, on va s'intéresser à l'influence de l'apprentissage actif sur les performances de notre méthode pour les vidéos. Pour rappel, l'apprentissage actif est une méthode qui va permettre de réinjecter des exemples négatifs dans la phase d'apprentissage afin d'améliorer les performances en réduisant le nombre de fausses détections. Et comme on a pu le voir pour les images fixes, l'apprentissage actif va certainement améliorer nos performances.

On va donc comparer les différentes méthodes de classification utilisées avec et sans apprentissage actif sur la base de données de 18 vidéos avec les métriques introduites

précédemment. Les résultats de cette expérimentation sont disponibles dans les tableaux 3.7 et 3.8. Afin d'identifier si la classification utilisée a bénéficié de l'étape d'apprentissage actif ou non, il suffit de regarder le numéro (par exemple le 1 dans *LBP $N$ 1*), ce numéro traduit le nombre de répétitions d'apprentissage actif appliqué à cette classification.

| Méthode                   | PDR  | MPT   | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT          |
|---------------------------|------|-------|------|-----------|--------|----------------------|-------------|
| <i>LBP<math>N</math>0</i> | 100% | 140ms | 3.2  | 12.41%    | 54.58% | 20.22%               | 7.2 (0.3s)  |
| <i>LBP<math>N</math>1</i> | 100% | 160ms | 1.2  | 23.89%    | 46.03% | 31.46%               | 13.5 (0.5s) |
| <i>LBP<math>N</math>2</i> | 100% | 162ms | 0.9  | 23.42%    | 35.08% | 28.08%               | 29.4 (1.2s) |

TABLE 3.7 – Résultats de la détection dans les vidéos avec l'apprentissage actif pour la classification utilisant les motifs binaires locaux (LBP)

| Méthode                    | PDR  | MPT  | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT          |
|----------------------------|------|------|------|-----------|--------|----------------------|-------------|
| <i>Haar<math>N</math>0</i> | 100% | 26ms | 1.4  | 23.29%    | 46.82% | 31.10%               | 17.5 (0.7s) |
| <i>Haar<math>N</math>1</i> | 100% | 23ms | 0.6  | 35.58%    | 42.29% | 38.65%               | 25.8 (1.0s) |

TABLE 3.8 – Résultats de la détection dans les vidéos avec l'apprentissage actif pour la classification utilisant les paramètres pseudo-Haar

Comme on peut le voir avec les tableaux 3.7 et 3.8, l'apprentissage actif n'impacte pas le PDR, on est donc capable d'alerter le médecin sur la présence du polype dans chacune des 18 vidéos. Côté performances :

- Avec la classification utilisant les motifs binaires locaux (tableau 3.7), quand on passe de la classification sans apprentissage actif à celle avec une étape de d'apprentissage actif, les résultats sont nettement améliorés : on a une augmentation du F<sub>1</sub> Score, qui se traduit par l'augmentation de la précision. Conséquence directe de cet effet, on réduit le nombre de MNFP, donc moins de fausses détections sont retournées au médecin. Cependant, avec une deuxième étape d'apprentissage actif, les performances commencent à se dégrader, même si le nombre de fausses détections diminue encore, il est donc intéressant de garder cette classification.
- Avec la classification utilisant les paramètres pseudo-Haar (tableau 3.8), on observe le même effet qu'avec les motifs binaires locaux pour une étape d'apprentissage actif, on a une augmentation du F<sub>1</sub> Score grâce à l'augmentation de la précision, et donc réduit le nombre de MNFP.

Dans les deux types de classification utilisés, on note une augmentation du temps de réaction. Cela est dû au fait que le classificateur devient de plus en plus sélectif, et donc va mettre un peu plus de temps pour trouver le polype après sa première apparition à l'image. Cependant, on reste de l'ordre de la seconde, ce qui est tout à fait acceptable.

### 3.2.3 Cohérence temporelle

On va maintenant s'intéresser ici à une des autres optimisations mise en place afin d'améliorer les performances de la méthode, et qui s'adapte uniquement aux vidéos : il s'agit de la cohérence temporelle.

Contrairement aux images fixes disponibles dans la base CVC-ColonDB [BSV12] qui ne se suivent pas forcément dans la prise de vue, la vidéo possède l'avantage d'avoir des images qui se suivent avec une variation minimale de l'information entre deux images successives. Il peut donc être intéressant ici d'utiliser cette faible variation entre la succession d'images afin d'améliorer nos performances.

On va donc mettre au point une cohérence temporelle basée sur la mise en mémoire des précédentes détections dans la vidéo afin de réduire le nombre de fausses détections. Cette cohérence temporelle vient donc se greffer à la méthode présentée précédemment qui, elle, ne change pas (qui utilise la fusion spatiale et si désiré, l'apprentissage actif). Le fonctionnement du nouvel algorithme est décrit dans le tableau 3.9.

- 
1. Une étape de prétraitement de l'image pour préparer la détection
  2. Une étape de détection cascade qui permet de trouver les régions d'intérêt sur l'image contenant de possibles polypes (avec ou sans apprentissage actif)
  3. Une étape de fusion spatiale des zones détectées afin de limiter le nombre de fausses détections et ainsi réduire le nombre de régions d'intérêt affichées à l'utilisateur
  4. Une étape de fusion temporelle des zones détectées pour l'image en cours avec les zones détectées des deux précédentes images
  5. Une étape d'affichage permettant de retourner le résultats de la détection à l'utilisateur
- 

TABLE 3.9 – Algorithme de détection des polypes dans les vidéos

#### Fonctionnement

On va alors s'intéresser à la mise au point de la cohérence temporelle. Elle se base sur l'utilisation des résultats de la détection sur 3 images successives. Pour obtenir le résultat de la détection dans l'image  $i$ , on va superposer sur l'image  $i$  les résultats des détections qui ont eu lieu dans les images  $i$ ,  $i - 1$  et  $i - 2$ . De cette superposition, on va fusionner ces différentes zones (de la même manière que la fusion spatiale) afin de réduire les fausses détections. Et c'est ce résultat qui est alors renvoyé à l'utilisateur. En voici un exemple sur la figure 3.4 qui illustre son fonctionnement : lors de la détection sur l'image  $i$ , on regarde si des détections ont eu lieu au même endroit ou a proximité sur les 2 précédentes images, si oui, dans ce cas la zone est considérée comme une bonne détection et est affichée à l'utilisateur. Sinon, la zone n'est pas affichée (mais est gardée en mémoire pour les

prochaines détections au cas où).

L'utilisation de 3 images se justifie par le fait d'utiliser un vote à la majorité pour fusionner les zones détectées, donc on a besoin d'un nombre impair d'images. On aurait pu utiliser plus d'images, mais cela risque d'augmenter le temps de réaction de la méthode, et comme on désire un temps de réaction des plus faibles, on décide alors de n'utiliser que 3 images, ce qui semble être un bon compromis.

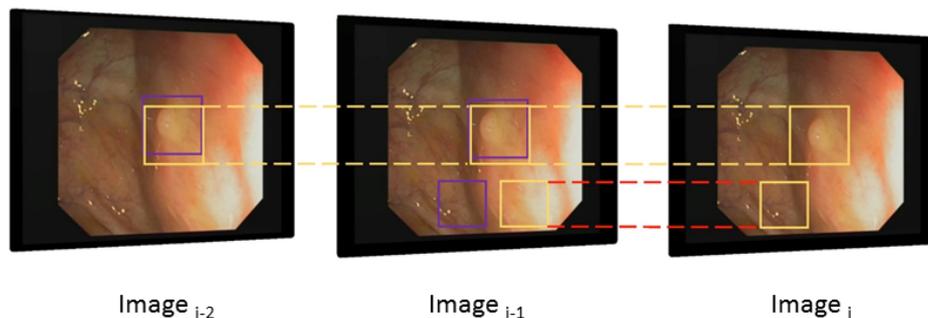


FIGURE 3.4 – Illustration de la cohérence temporelle

### Effet de la cohérence temporelle sur la détection

La cohérence temporelle va avoir pour but de lisser les résultats de détection, afin d'apporter une meilleure stabilité dans le résultat affiché à l'écran. Cela peut se voir indirectement sur les profils résultants des détections dans les vidéos (figure A.1 à A.18). La figure 3.5 (extraite de la figure A.5) illustre le lissage observé. Sur la figure 3.5.b, on peut observer le profil de la détection sans cohérence temporelle (1 si le polype est bien trouvé, 0 sinon) et sur la figure 3.5.c, on peut observer le profil de la détection avec la cohérence temporelle. On voit donc que la cohérence temporelle permet de "lisser" les résultats bruts, elle supprime les oscillations que l'on peut observer dû au changement de la présence du polype dans la vidéo. Cela permet donc d'avoir un meilleur rendu visuel pour le médecin en n'ayant pas des boîtes qui enchainent apparitions et disparitions.

### Performances sans apprentissage actif

Maintenant que la cohérence temporelle est mise en place, on peut donc s'intéresser à son influence sur les performances sans utiliser l'apprentissage actif pour le moment, c'est-à-dire qu'on cherche à voir simplement quels peuvent être les bénéfices d'utiliser une telle optimisation pour améliorer les résultats de base de notre méthode avec déjà un minimum de fausses détections. On va donc comparer les résultats de la détection avec et sans cohérence temporelle sur les deux types de paramètres utilisés pour l'apprentissage (les motifs locaux binaires (LBP) et les paramètres pseudo-Haar). Comme précédemment, ces tests

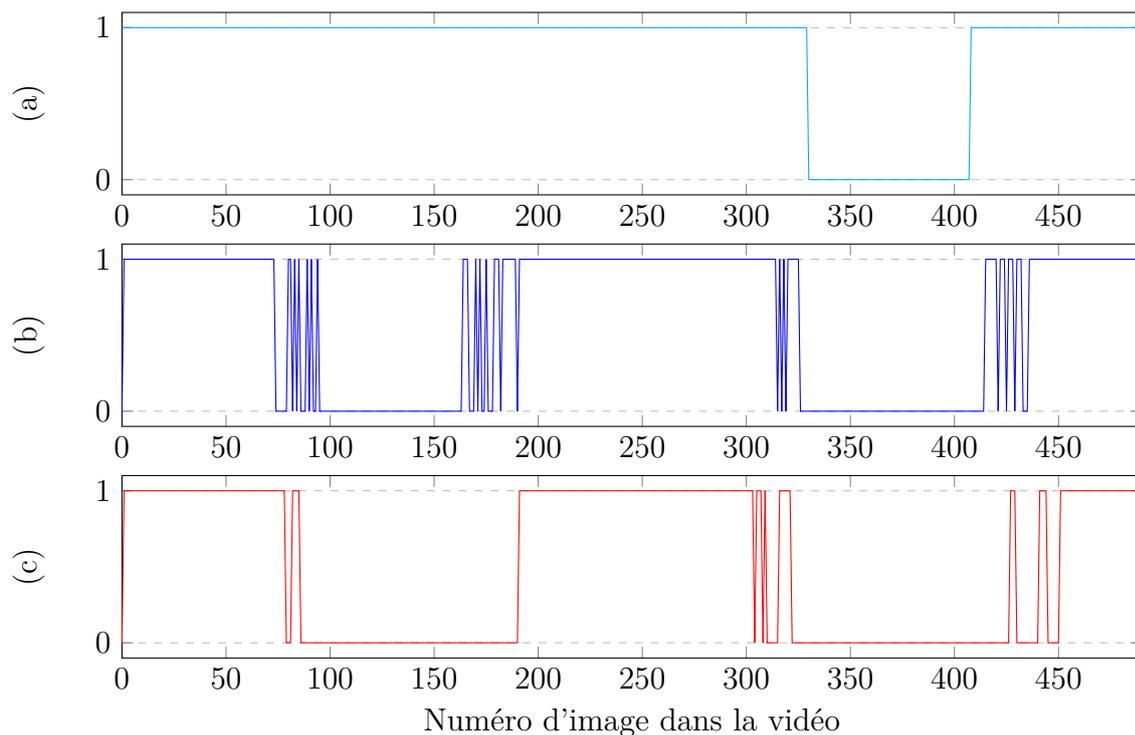


FIGURE 3.5 – Effet de la cohérence temporelle pour la vidéo 5 (490 images), (a) représente la vérité terrain, (b) représente les résultats de détection sans cohérence temporelle, (c) représente les résultats de détection avec cohérence temporelle

seront menés sur l'ensemble des 18 vidéos à notre disposition et évalués avec les métriques associées. Les résultats de ces différents tests sont consignés dans les tableaux 3.10 et 3.11.

| Cohérence | PDR  | MPT   | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT          |
|-----------|------|-------|------|-----------|--------|----------------------|-------------|
| Sans      | 100% | 140ms | 3.5  | 12.42%    | 54.65% | 20.24%               | 7.2 (0.3s)  |
| Avec      | 100% | 140ms | 1.9  | 16.25%    | 41.25% | 23.31%               | 35.0 (1.4s) |

TABLE 3.10 – Influence de la cohérence temporelle sur les résultats utilisant la classification basée sur les LBP sans apprentissage actif

| Cohérence | PDR  | MPT  | MNFP | Précision | Rappel | F <sub>1</sub> Score | RT          |
|-----------|------|------|------|-----------|--------|----------------------|-------------|
| Sans      | 100% | 26ms | 1.4  | 23.29%    | 48.82% | 31.10%               | 17.5 (0.7s) |
| Avec      | 100% | 26ms | 0.9  | 27.02%    | 39.61% | 32.12%               | 38.3 (1.5s) |

TABLE 3.11 – Influence de la cohérence temporelle sur les résultats utilisant la classification basées les paramètres pseudo-Haar sans apprentissage actif

A partir des résultats disponibles dans les tableaux 3.10 et 3.11, on s'aperçoit toujours qu'on est capable d'attirer l'attention du médecin sur le polype (PDR de 100%), ce qui

veut dire que la cohérence temporelle n'impacte pas le but premier de l'outil : alerter sur la présence du polype et sa position. Côté performance de détection, la cohérence temporelle permet de réduire le nombre moyen de fausses détections par image, qui pour les LBP descend en dessous de 2, et pour les paramètres pseudo-Haar descend en dessous de 1. Cela va donc impacter le  $F_1$  Score qui va augmenter, ce qui traduit une amélioration globale des performances de détection. Concernant le temps de calcul, on constate que dans dans les deux cas, le temps de calcul n'augmente pas de façon sensible pour être visible (les variations des temps de calcul sont inférieurs à la milliseconde). Par contre, la cohérence temporelle va augmenter le temps de réaction de notre méthode, ce qui est tout a fait logique car il faut plus d'images pour valider le résultat de la détection. La figure 3.6 permet de se rendre compte de l'influence de la cohérence temporelle sur les régions d'intérêts retournées au médecin.

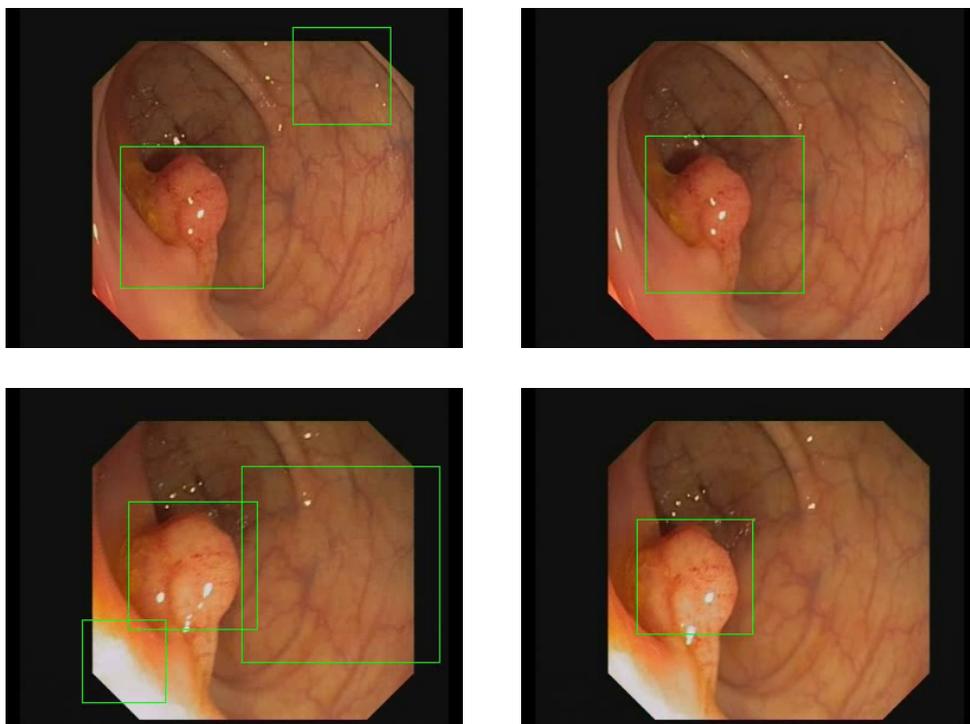


FIGURE 3.6 – Images illustrant la cohérence temporelle, à gauche des résultats de détection sans cohérence temporelle, à droite des résultats avec cohérence temporelle

### Temps de réaction

Une chose que l'on a pu constater sur la quasi totalité des vidéos, c'est que l'ajout de la cohérence temporelle a eu un effet sur le temps de réaction de notre outil. En effet, comme on peut le constater sur la tableau 3.12, la moyenne des temps de réaction des 18 vidéos passe du simple au double avec la cohérence. Mais comme il s'agit d'une moyenne, et que l'augmentation de la valeur de cette métrique n'est pas négligeable, on peut alors se demander si ce n'est pas une ou deux vidéos qui mettent à mal le temps de réaction.

Dans le tableau 3.13, on peut voir les temps de réaction pour chacune des vidéos sans la cohérence temporelle et dans le tableau 3.14, les temps de réaction avec la cohérence temporelle.

| Méthode        | Temps de réaction (images) |
|----------------|----------------------------|
| Sans cohérence | 17.5 (0.7s)                |
| Avec cohérence | 38.3 (1.5s)                |

TABLE 3.12 – Influence de la cohérence temporelle sur le temps de réaction avec la classification utilisant les paramètres pseudo-Haar sans apprentissage actif

|                              |      |      |      |      |      |      |      |      |      |
|------------------------------|------|------|------|------|------|------|------|------|------|
| Numéro de vidéo              | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
| Temps de réaction (secondes) | 0.6  | 3.44 | 0.08 | 0.16 | 0.08 | 0.08 | 0.08 | 3.44 | 0.32 |
| Numéro de vidéo              | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   |
| Temps de réaction (secondes) | 0.24 | 0.20 | 0.08 | 0.08 | 0.08 | 3.4  | 0.12 | 0.12 | 0    |

TABLE 3.13 – Temps de réaction pour chaque vidéo sans cohérence temporelle

|                              |      |      |      |      |      |      |      |      |      |
|------------------------------|------|------|------|------|------|------|------|------|------|
| Numéro de vidéo              | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
| Temps de réaction (secondes) | 1.28 | 3.4  | 0.08 | 1.2  | 0.08 | 0.08 | 6.28 | 8.52 | 1.68 |
| Numéro de vidéo              | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   |
| Temps de réaction (secondes) | 1.12 | 0.28 | 0.08 | 0.16 | 0.12 | 3.48 | 0.56 | 0.56 | 0    |

TABLE 3.14 – Temps de réaction pour chaque vidéo avec cohérence temporelle

On constate alors que le temps de réaction augmente de manière importante sur deux vidéos en particulier. Si on retire ces deux vidéos de l'équation, on obtient alors un temps de réaction de 19.94 images, soit environ 0.8 secondes, plutôt proche de la valeur obtenue sans cohérence temporelle (qui est de 14.19 images, soit environ 0.6 secondes). On observe seulement une différence de 0.2 secondes, contrairement aux 0.8 secondes observées au début, c'est donc bien ces deux vidéos qui détériorent les performances du temps de réaction. On pourrait alors conclure que la cohérence temporelle ne détériore que faiblement la détection des polypes si on fait abstraction des vidéos 7 et 8. Mais avant de conclure cela, il serait intéressant de regarder en détail les deux vidéos en question afin d'identifier la (ou les) cause(s) qui augmente le temps de réaction dans ces deux vidéos. On va donc traiter ces deux cas séparément par la suite.

### Cas de la vidéo 7

On s'intéresse au cas de la vidéo 7. On a pu constater que le temps de réaction sans cohérence temporelle est de 2 images, mais passe subitement à 157 avec la cohérence

temporelle. Quand on regarde le profil de cette vidéo (figure 3.8, on a 1 si le polype est présent dans l'image, 0 sinon), on s'aperçoit que le polype fait une série d'apparition et de disparition, et c'est ce qui semble être la cause de l'augmentation du temps de réaction. En effet, la cohérence temporelle a pour but de moyenner les détections, alors si peu de détections sont faites, alors le résultat de la cohérence temporelle donnera l'impression que l'algorithme n'a rien détecté. Ce qui explique l'augmentation du temps de réaction dans cette vidéo puisque ce type de structure peut effectivement apparaître.

Mais ce n'est pas le profil de la vidéo qui est à remettre en cause, mais la qualité de l'image, si on se rappelle le tableau 3.1 qui notait la difficulté à détecter le polype dans les vidéos, la vidéo 7 avait obtenu un score élevé, synonyme d'une forte difficulté à trouver le polype. Cette note élevée est due tout simplement à la qualité de la vidéo. Quand on regarde la vidéo, on peut s'apercevoir de la présence de liquide qui passe devant l'endoscope (figure 3.7), et qui introduit un certain flou, ce qui peut alors mettre en échec la méthode, et cela à plusieurs reprises. Comme la cohérence temporelle requiert 3 images pour fonctionner, si on n'est plus capable d'effectuer des détections sur certaines images, alors on va clairement augmenter notre temps de réaction. Et c'est donc la raison pour laquelle le temps de réaction augmente dans cette vidéo.

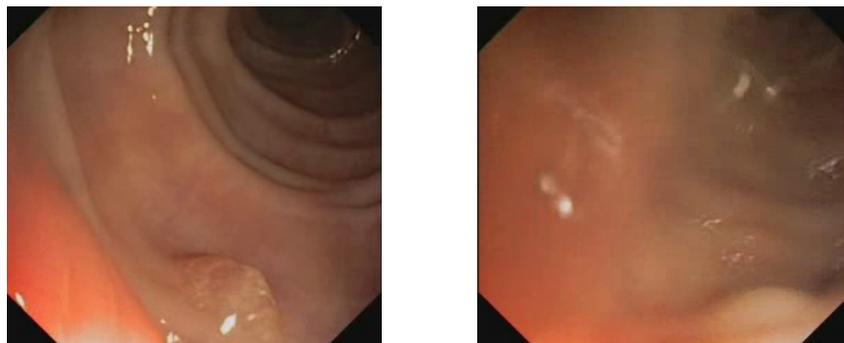


FIGURE 3.7 – Images extraites de la vidéo 7 (image nette à gauche et image floutée à cause d'un liquide à droite)

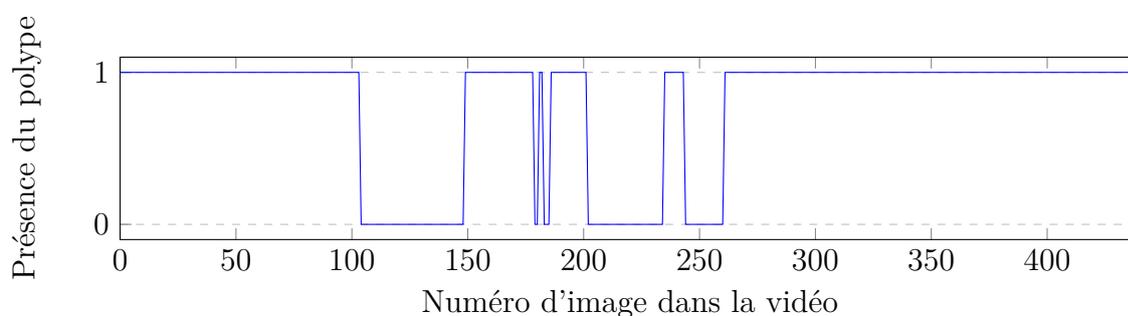


FIGURE 3.8 – Présence du polype dans la vidéo 7 (440 images)

### Cas de la vidéo 8

Dans le cas de la vidéo 8, le temps de réaction sans cohérence temporelle est de 86 images, mais passe subitement à 213. On constate ici que de base, l'algorithme met du temps à trouver le polype (car il est peut être difficile à trouver). Comme précédemment, on peut regarder le profil de la vidéo (figure 3.10) pour obtenir quelques indices, mais il n'y a pas de successions d'apparitions et de disparitions du polype dans la vidéo. Cependant, cela nous indique donc qu'il faut s'intéresser à l'image en elle même. Et comme on peut le constater en regardant la figure 3.9, le polype est difficilement différentiable de la paroi, d'où le score qu'avait obtenu cette vidéo dans le tableau 3.1. Cela explique donc d'une part pourquoi le temps de réaction était déjà élevé au départ, mais qu'en rajoutant la cohérence temporelle, on va donc indéniablement augmenter le temps de réaction.



FIGURE 3.9 – Images extraites de la vidéo 8 (image éloignée du polype à gauche et image proche du polype à droite)

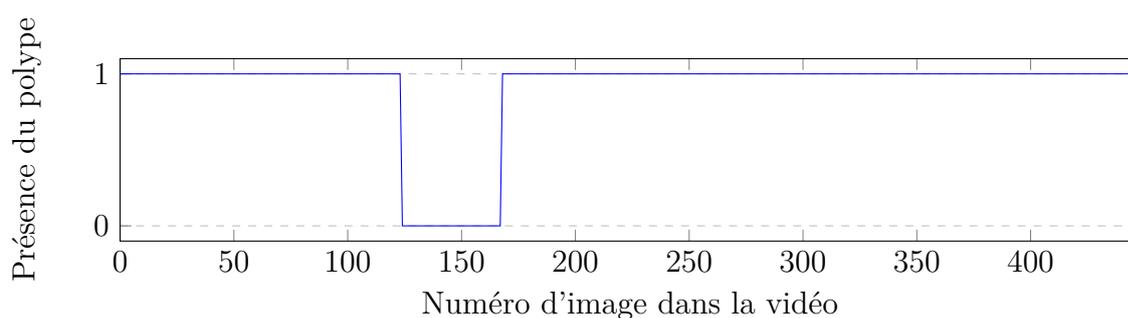


FIGURE 3.10 – Présence du polype dans la vidéo 8 (448 images)

### Conclusion sur la cohérence temporelle

La cohérence temporelle permet clairement de réduire le nombre moyen de fausses détections par image et permet donc de réduire les perturbations renvoyées au médecin.

Cependant, la cohérence temporelle rend l’outil moins réactif (au sens de sa rapidité à détecter un polype après sa première apparition dans à l’image, autrement dit le temps de réaction). De plus, le rappel diminue lorsque l’on applique la cohérence temporelle. Mais on va pu voir précédemment que l’apprentissage actif permettait justement d’augmenter le rappel de la méthode. On va donc maintenant s’intéresser à l’association de la cohérence temporelle avec l’apprentissage actif.

De plus, la cohérence va augmenter de façon significative le temps de réaction dans certains cas : lorsque le polype va enchaîner disparitions et apparitions, lorsque la qualité de la vidéo n’est pas bonne (qui peut être lié à la qualité de la préparation du côlon) ou lorsque le polype est difficilement différentiable de la paroi du côlon. Cependant, ces cas restent assez rares, et malgré cela, la cohérence temporelle permet quand même d’améliorer les performances globales de la méthode (qui s’aperçoit par l’augmentation du  $F_1$  Score).

### 3.2.4 Combinaison de l’apprentissage actif et de la cohérence temporelle

Tout comme pour le test précédent sur l’influence de la cohérence temporelle sur les performances de détection, on va maintenant effectuer un test pour analyser l’influence de l’apprentissage actif sur la méthode de détection dans les vidéos équipées de la cohérence temporelle dans le but d’améliorer les performances de la méthode. Ce test va donc aussi s’effectuer sur la base de données vidéo [ABSM<sup>+</sup>17b] avec les mêmes métriques. Les résultats de ce test sont disponibles dans le tableau 3.15

| Méthode   | PDR  | MPT   | MNFP | Précision | Rappel | $F_1$ Score | RT          |
|-----------|------|-------|------|-----------|--------|-------------|-------------|
| LBP $N0$  | 100% | 140ms | 1.9  | 16.25%    | 41.25% | 23.31%      | 35.0 (1.4s) |
| LBP $N1$  | 100% | 160ms | 1.1  | 27.11%    | 46.02% | 34.12%      | 43.7 (1.7s) |
| LBP $N2$  | 100% | 162ms | 0.7  | 29.88%    | 34.96% | 32.22%      | 45.9 (1.8s) |
| Haar $N0$ | 100% | 26ms  | 0.9  | 27.02%    | 39.61% | 32.12%      | 38.3 (1.5s) |
| Haar $N1$ | 100% | 23ms  | 0.6  | 39.14%    | 42.56% | 40.78%      | 27.3 (1.1s) |

TABLE 3.15 – Résultats de la détection dans les vidéos avec la cohérence temporelle et l’apprentissage actif

D’après le tableau 3.15, on constate que l’apprentissage actif associé à la cohérence temporelle permet d’obtenir les meilleures performances. Concernant la classification avec les LBP, on s’aperçoit que l’association de l’apprentissage actif avec la cohérence temporelle influe sur les performances de la classification (la précision augmente mais le rappel diminue) et le temps de calcul moyen par image augmente. Cela s’explique par le fait que l’apprentissage actif augmente la précision (comme vu précédemment) mais fait chuter le rappel, et il en est de même pour la cohérence temporelle. Donc à partir du moment où les deux sont associés, ces deux paramètres vont évoluer comme prévu, mais le  $F_1$  va

augmenter. Mais contrairement à la classification LBP, la classification utilisant les paramètres pseudo-Haar avec une étape d'apprentissage actif réunit les meilleurs résultats : à elle seule, elle possède un temps de calcul très faible par image, tout en ayant un  $F_1$  Score assez élevé. C'est donc cette classification là qu'il faudrait favoriser pour effectuer la détection des polypes lors des examens de coloscopie.

### Variations

Précédemment, on a pu voir le fonctionnement de la cohérence temporelle, où il s'agissait d'utiliser les résultats des deux précédentes détections pour valider ou non les détections dans l'image en cours. En se basant sur la figure 3.4, on utilisait les résultats de détection des images  $i$ ,  $i - 1$  et  $i - 2$  pour valider les résultats de l'image  $i$ . Mais on peut aussi envisager quelques variations de cette méthode de cohérence temporelle : utiliser le résultats de ces 3 détections pour valider la détection de l'image  $i - 1$  ou de l'image  $i - 2$ . Des tests ont été alors réalisés afin de regarder l'influence de ces variations. Ces tests ont été menés avec la classification utilisant les paramètres pseudo-Haar avec une étape d'apprentissage actif. Les résultats sont disponibles sur la figure 3.11.

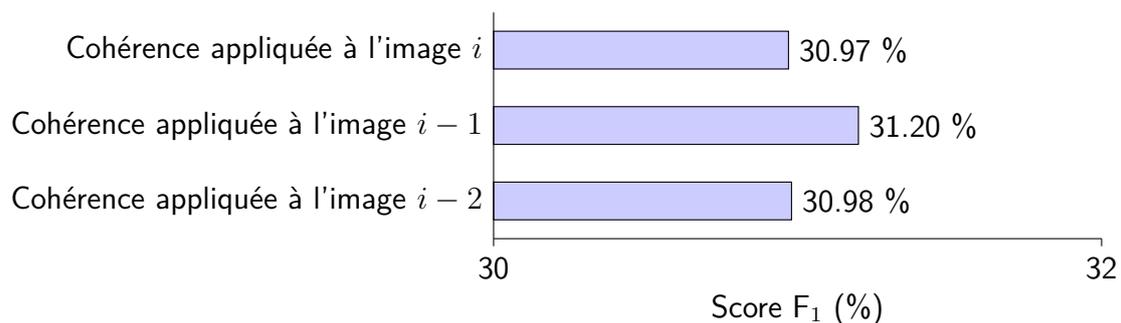


FIGURE 3.11 – Score  $F_1$  pour différentes combinaisons de cohérence temporelle

On constate donc que l'utilisation des résultats de la détection dans les images autour de l'image qu'on traite actuellement (c'est-à-dire utiliser les résultats des images  $i$  et  $i - 2$  pour valider les résultats de l'image  $i - 1$ ) permet d'améliorer sensiblement les performances de détection de la méthode. Cependant, l'utilisation de ces deux variations va rajouter un certain retard dans la chaîne de traitement : dans un cas, il faut attendre l'image d'après et les résultats de la détection, donc cela va rajouter un retard d'une image (et dans l'autre, cela rajoute un retard de deux images). Ainsi, au rythme de 25 images par secondes (rythme auquel arrive les images dans le flot vidéo), on va rajouter un retard entre 40 et 80 millisecondes. Cela va donc jouer évidemment sur le temps de réaction et l'augmenter, mais cette augmentation reste assez négligeable devant le temps de réaction de l'outil.

### 3.2.5 Combinaison des différents types de descripteurs d'apprentissage

Côté apprentissage, on a vu que les deux types de paramètres utilisés (motifs binaires locaux et paramètres pseudo-Haar) permettaient d'obtenir des performances différentes mais plutôt acceptables (si on fait abstraction du temps de calcul moyen, qui est dépendant de la puissance de la machine utilisée). On peut alors se poser la question de savoir si l'association de ces deux types d'apprentissage ne permettrait pas d'améliorer les performances globales de la méthode. On s'est donc intéressé à cette combinaison en associant la classification *LBP $N2$*  (classificateur utilisant les motifs binaires locaux avec une étape d'apprentissage actif) et *Haar $N1$*  (classificateur utilisant les paramètres pseudo-Haar avec une étape d'apprentissage actif). Le choix de ces classificateurs n'est pas dû au hasard, pour chaque type de paramètres utilisés lors de l'apprentissage, ce sont eux qui permettent d'obtenir les  $F_1$  Score les plus élevés. On mène donc les tests sur l'ensemble des 18 vidéos afin d'en calculer les performances. Les résultats de ce test sont disponibles dans le tableau 3.16.

| Méthode                    | PDR  | MPT   | MNFP | Précision | Rappel | $F_1$ Score | RT          |
|----------------------------|------|-------|------|-----------|--------|-------------|-------------|
| <i>LBP<math>N2</math></i>  | 100% | 162ms | 0.7  | 29.88%    | 34.96% | 32.22%      | 45.9 (1.8s) |
| <i>Haar<math>N1</math></i> | 100% | 21ms  | 0.6  | 39.14%    | 42.56% | 40.78%      | 27.3 (1.1s) |
| Combinaison                | 100% | 185ms | 1.1  | 30.39%    | 52.40% | 38.47%      | 15.0 (0.6s) |

TABLE 3.16 – Comparaison entre classifications classiques et combinaison de classifications (*LBP $N2$*  & *Haar $N1$* )

On voit donc sur le tableau 3.16 que la combinaison des deux types de classification permet toujours d'alerter le médecin sur la présence du polype dans l'ensemble des 18 vidéos, mais qu'en plus, le Rappel augmente, ce qui signifie qu'on est capable de détecter plus de polype qu'auparavant mais de manière moins précise. On peut aussi voir que le temps de réaction diminue fortement, ce qui induit une meilleure rapidité dans la détection après la première apparition du polype à l'image. En somme, la combinaison des deux classifications est une piste à explorer car elle permet d'obtenir de meilleurs résultats. Cependant, le temps de calcul reste trop élevé pour faire du temps réel (on hérite du temps de calcul de la classification utilisant les motifs binaires locaux), et donc cela ne nous convient pas pour le moment.

### 3.2.6 Prétraitement de l'image

On a pu voir plusieurs pistes qui ont permis d'améliorer les performances de la méthode, ces méthodes se sont principalement concentrées sur l'amélioration de l'apprentissage actif et la réduction des différentes régions d'intérêt détectées en appliquant une fusion spatiale et une cohérence temporelle. Cependant, une autre piste peut être explorée pour améliorer les performances. Il s'agit d'appliquer un prétraitement à l'image avant d'effectuer la moindre détection. On va donc s'intéresser dans cette partie aux différentes

possibilités de prétraitement de l'image et leurs influences sur les performances de la méthode.

### Influence du retrait d'artefacts

Une première piste de prétraitement vient lorsqu'on regarde les images issues de vidéo-coloscopies. On s'aperçoit qu'une bonne partie de l'image ne contient pas d'information permettant la détection des polypes. Il s'agit de tout le contour de l'image en noir (voir figure 3.12). Il peut être alors intéressant de comparer les performances de détection entre des images normales et des images détournées (voir figure 3.12). La comparaison est alors effectuée pour un seul type de classification pour voir la différence (la classification utilisant les paramètres pseudo-Haar est utilisée avec une étape d'apprentissage actif). Les résultats sont disponibles dans le tableau 3.17.

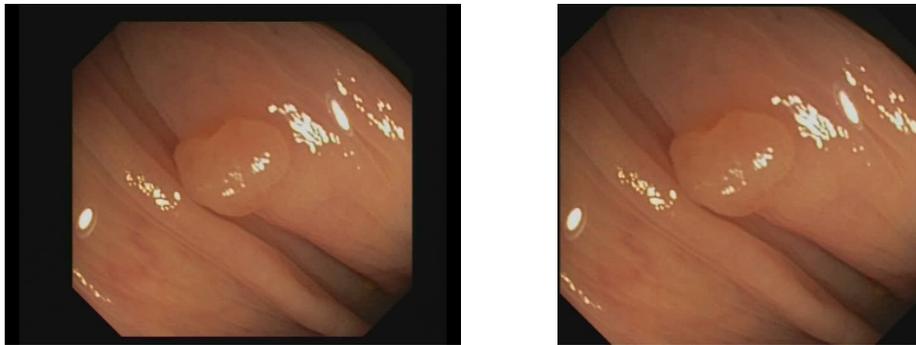


FIGURE 3.12 – Image avec les bordures (à gauche) et image sans (à droite)

| Images        | TP   | FP   | TN  | FN   | Précision | Rappel | F <sub>1</sub> Score | PDR  |
|---------------|------|------|-----|------|-----------|--------|----------------------|------|
| Originales    | 4267 | 6634 | 702 | 5758 | 39.14%    | 42.56% | 40.78%               | 100% |
| Sans contours | 4496 | 5338 | 880 | 5529 | 45.72%    | 44.85% | 45.28%               | 100% |

TABLE 3.17 – Comparaisons des performances de détection entre images normales et images détournées

On constate alors qu'en retirant les bordures noires, on améliore globalement les performances de la détection. Certes, le nombre de bonnes détections diminue, mais le nombre de fausses détections diminue drastiquement, ce qui est un avantage car cela retourne moins d'informations erronées au médecin, ce qui peut améliorer les performances de l'examen.

Cela peut s'expliquer de la manière suivante : le retrait des bordures de l'image retire de l'information, mais pas une information si importante que ça. En retirant donc cette information, certes on perd en performance, mais on permet à l'algorithme de mieux se concentrer sur les zones à détecter et permet alors de diminuer le nombre de fausses détections.

| Images        | Temps de calcul |
|---------------|-----------------|
| Originales    | 23 ms           |
| Sans contours | 16 ms           |

TABLE 3.18 – Comparaisons des temps de traitement entre images normales et images détournées

Mais de plus, en retirant ces bordures, on accélère le temps de traitement (figure 3.18) car l'image à traiter diminue en résolution, sans perdre d'information utile. Donc détourner les images de leurs contours noirs présente de nombreux avantages, et peut être une des solutions simples afin d'améliorer les performances de notre outil.

### Influence du retrait de la lumière spéculaire

La seconde piste s'obtient aussi en regardant les images issues de coloscopies. On peut apercevoir sur plusieurs images la réflexion de la lumière de l'endoscope sur la paroi du côlon : il s'agit de la lumière spéculaire (voir la figure 3.13). On a pu voir que sur plusieurs images renvoyées par la méthode (utilisée pour la détection des polypes) que des régions d'intérêt se trouvent être des zones contenant de la lumière spéculaire et pas des polypes. On peut donc essayer de supprimer cette lumière dans l'optique d'améliorer les résultats de détection. Pour ce faire, on utilise une méthode mise au point par J. Bernal [BSV13] qui va permettre de détecter les zones contenant la lumière spéculaire et repeindre ces zones avec des pixels voisins afin de réduire leurs impacts sur la détection (voir la figure 3.13). La comparaison va alors être effectuée avec un seul type de classification pour voir la différence (la classification utilisant les paramètres pseudo-Haar est utilisée avec une étape d'apprentissage actif). Les résultats sont disponibles dans le tableau 3.19.

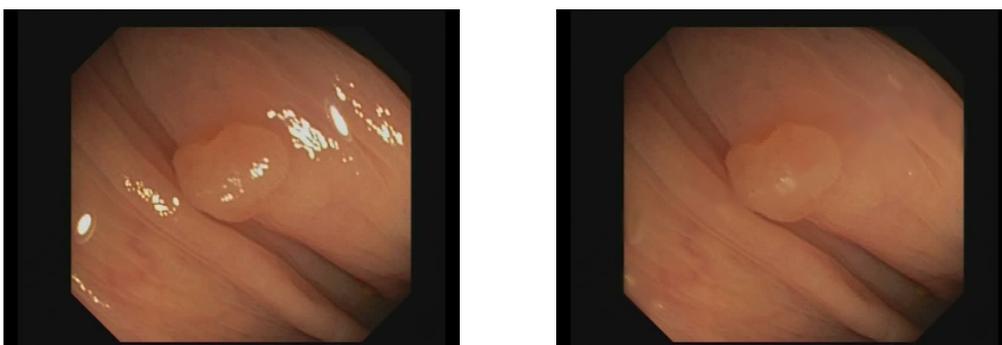


FIGURE 3.13 – Image avec lumière spéculaire (à gauche) et image inpaintée (à droite)

Selon le tableau 3.19, on constate alors une très légère détérioration des performances, en effet la Précision augmente mais le Rappel diminue, ce qui impacte évidemment le  $F_1$  Score, dont la variation reste minime. Dans la réalité, on se rend compte qu'on diminue le nombre de bonnes détections mais on diminue surtout le nombre de fausses détections,

| Images                  | TP   | FP   | TN  | FN   | Précision | Rappel | F <sub>1</sub> Score | PDR  |
|-------------------------|------|------|-----|------|-----------|--------|----------------------|------|
| Originales              | 4267 | 6634 | 702 | 5758 | 39.14%    | 42.56% | 40.78%               | 100% |
| Sans lumière spéculaire | 4342 | 6934 | 669 | 5683 | 38.51%    | 43.31% | 40.77%               | 100% |

TABLE 3.19 – Comparaisons des performances de détection entre images normales et images sans lumière spéculaire

| Images        | Temps de calcul |
|---------------|-----------------|
| Originales    | 23 ms           |
| Sans contours | 20 ms           |

TABLE 3.20 – Comparaisons des temps de traitement entre images normales et images détournées

cela traduit donc toujours la capacité à détecter les polypes tout en fournissant un faible nombre de fausses alertes, ce qui perturbe moins le médecin. On est alors capable de mieux attirer l'attention du médecin sur la présence du polype, sachant qu'on conserve toujours un PDR de 100%, indiquant qu'on alerte le médecin sur la présence de tous les polypes. L'obtention de tels résultats s'explique par le fait que de base, la méthode détectait certaines zones de lumière spéculaire comme étant des polypes, donc en retirant cette lumière spéculaire, on retire les fausses détections qui y sont liées, et on perturbe moins la détection. Il s'en suit alors une légère modification des performances, comme observée. De plus, d'après le tableau 3.20, on constate que le temps de calcul diminue, ce qui montre un second intérêt d'utiliser une telle optimisation.

### Influence de la combinaison des deux phases de prétraitement

Précédemment, on a constaté que retirer les bordures permettait de diminuer de façon conséquente les fausses détections mais diminuait aussi les bonnes détections. En retirant la lumière spéculaire, on s'apercevait qu'on augmentait le nombre de bonnes détections mais aussi, on a eu quelques fausses détections. Naturellement, on peut se poser la question de combiner les deux. Pour cela, on utilise donc des images recadrées et retouchées (figure 3.14). On pourrait donc s'attendre à une chute de fausses détections, mais aussi une baisse des bonnes détections, mais moins conséquente qu'avec uniquement le retrait des bordures. Les essais ont été menés avec la classification utilisant les paramètres pseudo-Haar (avec une étape d'apprentissage actif). Les résultats sont disponibles dans le tableau 3.21.

On constate alors qu'on améliore bien les performances globales de la méthode (F<sub>1</sub> Score). On peut donc envisager d'intégrer une étape de prétraitement dans la chaîne de détection afin de retirer les bordures et la lumière spéculaire afin de limiter grandement le nombre de fausses détections, et donc de moins perturber le médecin durant son examen.



FIGURE 3.14 – Image sans bordure et sans lumière spéculaire

| Images                                   | TP   | FP   | TN  | FN   | Précision | Rappel | F <sub>1</sub> Score | PDR  |
|--|------|------|-----|------|-----------|--------|----------------------|------|
| Originales                               | 4267 | 6634 | 702 | 5758 | 39.14%    | 42.56% | 40.78%               | 100% |
| Sans contours                            | 4496 | 5338 | 880 | 5529 | 45.72%    | 44.85% | 45.28%               | 100% |
| Sans lumière spéculaire                  | 4342 | 6934 | 669 | 5683 | 38.51%    | 43.31% | 40.77%               | 100% |
| Sans contours et sans lumière spéculaire | 4630 | 5438 | 915 | 5395 | 45.98%    | 46.18% | 46.09%               | 100% |

TABLE 3.21 – Comparaisons des performances de détection entre images normales et images détournées

### 3.3 Conclusion sur la détection dans les vidéos

Dans cette partie, on a pu s'intéresser à la détection et la localisation des polypes dans les vidéos de coloscopies, vidéos qui se rapprochent de la réalité d'un examen, et qui vont permettre de prouver l'utilité d'un tel outil d'aide à la détection. On a donc pu voir dans cette partie quelle était la méthode retenue pour effectuer la détection des polypes dans les vidéos issues de coloscopies, méthode basée sur l'apprentissage machine pur utilisant soit les paramètres à motif binaire local ou les paramètres pseudo-Haar, et ainsi que ses limitations en terme de performances sur les vidéos grâce à l'introduction de nouvelles métriques dédiées et cliniquement utilisables (c'est-à-dire facilement interprétables par le médecin). Pour palier à ces limitations, plusieurs améliorations ont été développées pour améliorer les performances de notre outil.

La première amélioration est tout simplement la mise en place d'une fusion spatiale des régions d'intérêt détectées afin de réduire le nombre de régions d'intérêt inutiles et de réduire le nombre de fausses détections. On a ainsi vu que, quelque soit les paramètres d'apprentissage utilisés, les performances étaient améliorées. De plus, cette amélioration permettant un gain non négligeable en performances, il a été intégré et utilisé avec toutes les autres améliorations.

La seconde amélioration, qui avait été présentée dans la partie sur la détection dans les images fixes, est l'utilisation de l'apprentissage actif pour la détection dans les vidéos.

Encore une fois, l'apprentissage actif a montré ses avantages et a amélioré nettement les performances de notre méthode.

La troisième amélioration mise au point s'applique très bien au cas de la détection dans la vidéo. Il s'agit de la cohérence temporelle, cette amélioration permet d'obtenir de meilleurs résultats sur une image de la vidéo en se basant sur les détections obtenues dans les images précédentes, il s'agit donc d'une sorte de mémoire afin de valider un futur résultat. On a alors pu constater que la cohérence temporelle permettait d'améliorer les performances de notre outil.

Par la suite, on a combiné la seconde amélioration (l'apprentissage actif) avec la troisième (la cohérence temporelle). On a ainsi pu voir que cette combinaison permettait bel et bien d'améliorer les performances de notre outil et donc que ces deux améliorations ne sont pas néfastes l'une pour l'autre.

Cette association des trois améliorations montre de très bonnes performances (en particulier pour la classification utilisant les paramètres pseudo-Haar) et promet de bonnes choses pour un essai clinique. Mieux, l'ajout ces différentes améliorations n'impacte pas le temps de calcul. Il est tout à fait négligeable car les modifications des temps de calcul engendrées sont inférieures à la milliseconde.

On peut alors remarquer que toutes ces améliorations s'intéressent soit à l'étape de détection (l'apprentissage actif) ou à l'étape de post-traitement de la détection (fusion spatiale et cohérence temporelle). On s'est alors posé la question de savoir si il était possible de mettre au point des améliorations qui allaient se placer avant la détection, c'est-à-dire effectuer un pré-traitement sur l'image. On a donc relevé deux nouvelles améliorations possibles qui pourraient permettre d'avoir de meilleurs résultats.

La quatrième amélioration est le retrait de l'information inutile de l'image pour concentrer les recherches dans ce qui nous intéresse. On a alors retiré les bordures de chaque image car elles ne contiennent logiquement pas de polype et donc ne sont pas très utiles. On observe donc une amélioration des performances concernant la détection ainsi qu'une réduction du temps de calcul pour traiter une image. Cela veut dire qu'on peut consacrer un peu plus de temps à d'autres étapes de pré-traitement de l'image.

La cinquième amélioration est le retrait de la lumière spéculaire. Au cours des différents tests, la lumière spéculaire générait de nombreuses fausses détections, la question de la retirer fut judicieuse. En effet, le remplacement de cette lumière par les zones voisines permet d'obtenir des performances intéressante. Cependant, cette étape va rajouter nécessairement du temps de calcul, qui est au maximum de 19 millisecondes (rajouté aux 23 millisecondes, la contrainte de temps réel n'est plus atteinte).

Cependant, en combinant ces deux améliorations avec toutes les autres présentées, on améliore nettement les performances de notre outil, ce qui permet d'obtenir une méthode beaucoup plus robuste pour la détection des polypes les vidéos. De plus, avec toutes ces améliorations, l'outil est toujours capable de fonctionner presque en temps réel (en

utilisant uniquement les paramètres pseudo-Haar pour la détection). Les vidéos se rapprochant de cas réels, et avec notre méthode éprouvée, on peut donc envisager d'effectuer des essais cliniques en conditions réelles afin de se comparer au médecin et encore mieux prouver l'intérêt de notre outil de détection des polypes en temps réel.

Le tableau 3.22 résume les gains en performance de la méthode basée sur la classification utilisant les paramètres pseudo-Haar, classification, qui à l'heure actuelle, est la seule qui est capable de fonctionner en temps réel dans nos conditions expérimentales.

La tableau 3.22 montre clairement que les améliorations apportées à l'outil permettent d'obtenir de meilleurs résultats qu'à l'origine (on a un gain allant jusqu'à 29.82 points)

| Amélioration                        | Gain en performances (points) |
|-------------------------------------|-------------------------------|
| Aucune                              | +0                            |
| Fusion spatiale (1)                 | +14.83                        |
| Apprentissage actif (2) + 1         | +22.38                        |
| Cohérence temporelle (3) + 1        | +15.85                        |
| 1 + 2 + 3                           | +24.51                        |
| Retrait d'artefacts (4) + 1 + 2 + 3 | +29.01                        |
| Inpainting (5) + 1 + 2 + 3          | +24.50                        |
| 1 + 2 + 3 + 4 + 5                   | +29.82                        |

TABLE 3.22 – Résumé des gains en performances suivant les optimisations (Gains relatifs basés sur le  $F_1$  Score et sur la classification utilisant les paramètres pseudo-Haar)

On a donc mis au point une méthode permettant de localiser les polypes dans les vidéos de coloscopies, vidéos qui se rapprochent des cas réels. La méthode obtenue possède de très bonnes performances de détections ( $F_1$  Score de 40.78% avec pour la classification Haar $N1$  avec cohérence temporelle) et alerte le médecin sur la présence de polypes dans chacune des vidéos (PDR de 100%). La méthode fonctionne aussi en temps réel (23 millisecondes pour traiter une image). On a donc une méthode qui possède de bonnes performances, fonctionne en temps réel, embarquable et qui est cliniquement utilisable. On a donc mis au point une méthode répondant à nos contraintes. Par rapport à l'état de l'art, il semble alors qu'on soit les seuls à proposer ce genre d'outil, qui sera bientôt testé en routine clinique.

### 3.4 Évaluation de la robustesse

On a vu précédemment que notre méthode de détection offrait de bonnes performances. Mais on a décidé de tester notre méthode sur une nouvelle base de données contenant à nouveau 18 vidéos [BHM<sup>+</sup>18] afin de voir si les performances restent constantes ou changent grandement. Contrairement à la première base de données [ABSM<sup>+</sup>17b] qui contenait 11 954 images pour 18 vidéos, la nouvelle base de données [BHM<sup>+</sup>18] contient

elle 29 657 images pour 18 vidéos. Parmi ces 29 657 images, 21 813 contiennent un polype, soit 73.55% des images. On va ainsi mesurer les performances de notre méthode sur ces nouvelles méthodes en utilisant une partie des métriques précédentes, qui sont la Précision, le Rappel, le  $F_1$  Score, le temps de réaction (RT, en nombre d'images) et le taux de polypes détectés (PDR). Les résultats de ce test sont disponibles dans le tableau 3.23. Les tests menés se sont concentrés sur les classifications bénéficiant de la fusion spatiale, de la cohérence temporelle et de l'apprentissage actif (et ne bénéficiant d'aucune étape de prétraitement).

| Méthode   | Précision | Rappel | $F_1$ Score | RT          | PDR   |
|-----------|-----------|--------|-------------|-------------|-------|
| LBP $N0$  | 13.52%    | 39.11% | 20.09%      | 42.4 (1.7s) | 100%  |
| LBP $N1$  | 32.22%    | 36.98% | 34.44%      | 49.7 (2.0s) | 100%  |
| LBP $N2$  | 34.32%    | 37.14% | 35.67%      | 40.7 (1.6s) | 100%  |
| Haar $N0$ | 24.49%    | 35.93% | 29.13%      | 35.9 (1.4s) | 94.4% |
| Haar $N1$ | 33.28%    | 37.26% | 35.16%      | 25.8 (1.0s) | 100%  |

TABLE 3.23 – Résultats de la détection sur une nouvelle base de données vidéos

D'après le tableau 3.23, on constate qu'à part pour la classification Haar $N0$ , on est capable d'alerter l'oeil du médecin sur la position du polype dans l'image (PDR de 100%). Concernant le cas de la classification Haar $N0$ , ceci est dû à la difficulté à trouver le polype dans cette vidéo (le PDR de 94.4% est dû au rapport 17 sur 18). Concernant les performances, on s'aperçoit, comme précédemment que c'est la classification Haar $N1$  qui permet d'obtenir les meilleurs résultats en termes de performances de détection et localisation, mais aussi en temps de réaction, on se rapproche de la seconde, ce qui est tout à fait acceptable. Cela confirme aussi que c'est la classification la plus judicieuse à utiliser pour détecter les polypes.

Mais plus encore, quand on compare les résultats du test sur la nouvelle base de données par rapport aux anciens résultats 3.24, on s'aperçoit que les performances gardent une certaine stabilité. En effet, on constate (en mettant à part la classification Haar $N0$ ), que les résultats sont proches entre les deux bases de données. Cela montre donc la robustesse de notre outil à traiter de nouveaux cas, et sa capacité à détecter et localiser les polypes en toutes circonstances si on utilise la classification Haar $N1$  qui permet de conjuguer bonnes performances et temps de calcul faible (qui est de 20 millisecondes sur la base [BHM<sup>+</sup>18] et de 23 millisecondes sur la base [ABSM<sup>+</sup>17b]) afin de fonctionner en temps réel.

### 3.5 Détection de la présence des polypes dans les vidéos

Jusque là, la méthode est utilisée pour localiser les polypes dans les vidéos afin d'attirer l'oeil du médecin sur le polype en lui-même. Sur l'aspect localisation, la méthode

| Méthode                                | Précision | Rappel | F <sub>1</sub> Score | RT          | PDR   |
|--|-----------|--------|----------------------|-------------|-------|
| LBP <i>N0</i> [BHM <sup>+</sup> 18]    | 13.52%    | 39.11% | 20.09%               | 42.4 (1.7s) | 100%  |
| LBP <i>N0</i> [ABSM <sup>+</sup> 17b]  | 16.25%    | 41.25% | 23.31%               | 35.0 (1.4s) | 100%  |
| LBP <i>N1</i> [BHM <sup>+</sup> 18]    | 32.22%    | 36.98% | 34.44%               | 49.7 (2.0s) | 100%  |
| LBP <i>N1</i> [ABSM <sup>+</sup> 17b]  | 27.11%    | 46.02% | 34.12%               | 43.7 (1.7s) | 100%  |
| LBP <i>N2</i> [BHM <sup>+</sup> 18]    | 34.32%    | 37.14% | 35.67%               | 40.7 (1.6s) | 100%  |
| LBP <i>N2</i> [ABSM <sup>+</sup> 17b]  | 29.88%    | 34.96% | 32.22%               | 45.9 (1.8s) | 100%  |
| Haar <i>N0</i> [BHM <sup>+</sup> 18]   | 24.49%    | 35.93% | 29.13%               | 35.9 (1.4s) | 94.4% |
| Haar <i>N0</i> [ABSM <sup>+</sup> 17b] | 27.02%    | 39.61% | 32.12%               | 38.3 (1.5s) | 100%  |
| Haar <i>N1</i> [BHM <sup>+</sup> 18]   | 33.28%    | 37.26% | 35.16%               | 25.8 (1.0s) | 100%  |
| Haar <i>N1</i> [ABSM <sup>+</sup> 17b] | 39.14%    | 42.56% | 40.78%               | 27.3 (1.1s) | 100%  |

TABLE 3.24 – Comparaison des résultats de la détection sur les deux bases de données vidéos ([ABSM<sup>+</sup>17b] et [BHM<sup>+</sup>18])

montrait de bonnes performances. Mais on peut s'intéresser aussi aux performances de notre outil, non pas pour la localisation des polypes, mais pour la détection de la présence des polypes dans les vidéos, c'est-à-dire signaler au médecin le fait qu'un polype est visible à l'écran, mais sans le localiser.

Dans le cadre de notre collaboration avec l'institut Computer Vision Center de Barcelone, une comparaison entre notre méthode (nommée ALCAPOD pour l'occasion, Active Learning for Computer Aided POLyp Detection), la méthode développée par J. Bernal (nommée WM-DOVA [BSFE<sup>+</sup>15]) et les réseaux de neurones convolutifs (CNN) a été effectuée (présentée dans [BHM<sup>+</sup>18]) concernant la détection de la présence des polypes dans les vidéos. En effet, l'utilisation des réseaux de neurones est de plus en plus présente pour la détection d'objets ou de personnes et devient une des méthodes de référence. Pour mesurer les performances, on utilise toujours les mêmes métriques présentées jusqu'ici : la Précision, le Rappel, le F<sub>1</sub> Score, le temps de réaction (RT) et le temps de calcul moyen pour traiter une image (MPT). Les méthodes ont été testées sur la base de données présentées dans [BHM<sup>+</sup>18] qui contient 18 vidéos pour un total de 29 657 images. Les résultats sont disponibles dans le tableau 3.25.

D'après le tableau 3.25, on peut voir que notre outil, en termes de performance pour la détection de la présence des polypes dans les vidéos, se place bien par rapport aux méthodes utilisant les réseaux de neurones. On peut aussi voir que par rapport aux réseaux de neurones, on obtient une précision plus élevée au détriment du rappel et du temps de réaction. Cependant, on est capable de traiter les images beaucoup plus rapidement, si on compare notre méthode par rapport à celle utilisant les "CNN (FT) + C", on a un gain de 14 millisecondes sur le temps de calcul. On est alors parfaitement capable de détecter la présence de polypes dans les vidéos en temps réel avec de très bonnes performances. De plus, contrairement aux réseaux de neurones convolutifs, notre outil est embarquable, ce qui est un argument non négligeable pour l'utilisation clinique. Notre outil serait donc ici le seul à être retenu pour cette application particulière qu'est la détection de la présence de polypes colorectaux dans les vidéos.

| Méthode         | Précision    | Rappel       | F <sub>1</sub> Score | RT           | MPT           |
|-----------------|--------------|--------------|----------------------|--------------|---------------|
| WM-DOVA         | 80.2%        | 64.7%        | 71.6%                | 0.53s        | 10.50s        |
| ALCAPOD         | <b>81.4%</b> | 79.3%        | 80.3%                | 0.12s        | <b>0.020s</b> |
| CNN (FT)        | 79.6%        | 86.9%        | 83.1%                | 0.06s        | 0.038s        |
| CNN (FT) + C    | 80.0%        | <b>91.4%</b> | <b>85.3%</b>         | <b>0.05s</b> | 0.034s        |
| CNN (FT,DA)     | 80.4%        | 85.6%        | 82.9%                | 0.06s        | 0.038s        |
| CNN (FT,DA) + C | 80.5%        | 88.9%        | 84.5%                | <b>0.05s</b> | 0.028s        |
| CNN (S)         | 75.6%        | 74.7%        | 75.1%                | 0.23s        | 0.038s        |
| CNN (S) + C     | 75.5%        | 79.9%        | 77.6%                | 0.21s        | 0.033s        |
| CNN (S,DA)      | 74.1%        | 86.3%        | 79.7%                | 0.07s        | 0.038s        |
| CNN (S,DA) + C  | 73.4%        | 89.2%        | 80.5%                | 0.06s        | 0.033s        |

TABLE 3.25 – Comparaison des résultats de la détection de la présence des polypes dans les vidéos de la base [BHM<sup>+</sup>18]. FT : réglage fin, S : à partir de rien, DA : augmentation des données, C : compression

# Chapitre 4

## Conclusion et Discussions

### 4.1 Conclusion

On a donc pu voir que le cancer colorectal était une des causes principales de décès par cancer dans le monde, et surtout dans les pays développés. Le seul moyen pour lutter contre cette maladie est d'effectuer des examens de dépistages précoces, en pratiquant l'examen de référence qui se trouve être la coloscopie. Au cours de cette coloscopie, le médecin cherche les lésions à l'origine de ce cancer qui sont les polypes. Cependant, il peut arriver que le médecin rate des polypes durant l'examen, et cela peut monter jusqu'à 26%.

Afin de réduire ce chiffre, et donc augmenter le taux de détection des polypes, de nouveaux outils ont été développés comme de nouvelles méthodes de coloscopies (tel que le passage à la haute définition ou l'exploration d'images dans le spectre du proche visible). Cependant, ces nouveaux outils restent marginaux et très peu utilisés (sauf la haute définition qui est devenue le nouveau standard), et rien ne détrône la coloscopie classique qui reste ainsi l'examen de référence. C'est alors que sont apparues des méthodes de détection assistée par ordinateur. Ces nombreuses méthodes permettent d'analyser avec plus ou moins de réussite les images issues de coloscopies afin de trouver de manière automatique les polypes.

C'est ainsi dans ce contexte de méthodes de détection assistée par ordinateur qu'on se propose de mettre au point une nouvelle méthode permettant de détecter les polypes de façon automatique dans les images et vidéos de coloscopie. Cette méthode devant fonctionner en parallèle de l'examen pour donner un retour rapide au médecin, elle doit donc avoir plusieurs contraintes : elle devra donc fonctionner en temps réel, avoir de bonnes performances de détection (c'est-à-dire capable de détecter un grand nombre de polypes avec un minimum de fausses détections) et être embarquable dans les salles d'examens (autrement dit, proposer un système se rapprochant du plug'n'play facilement connectable et adaptable au colonnes d'endoscopie).

Après un bilan des différentes méthodes d'aide à la détection existantes, une première méthode, se plaçant dans la famille des méthodes hybrides, a été mise au point en se basant sur la détection de formes géométriques dans l'image (la recherche de formes circulaires en utilisant la transformée de Hough) qui retourne alors plusieurs régions d'intérêt dans

l'image. Cependant, on constate un fort taux de fausses détections. Afin de discriminer ces régions pour localiser le polype (et donc réduire le nombre de fausses détections), on effectue une analyse de texture (utilisant les paramètres d'Haralick) de ces régions, puis par comparaison avec une étape d'apprentissage machine (basée sur la technique de boosting, en particulier la méthode de Viola et Jones), on est plus ou moins capable de déterminer si la région détectée est un polype ou non. L'avantage principal de cette méthode est sa simplicité de calcul, qui en fait une méthode facilement embarquable. Malgré de bons résultats en terme de détection, la méthode se montre limitée : le temps de calcul dépend énormément de la taille de l'image en entrée, et plus l'image est grande, plus le temps de calcul augmente. Et plus encore, lors de tests sur des vidéos prises en haute définition, la méthode mettait plus d'une seconde à traiter une image, donc bien loin de tenir le temps réel (qui se doit d'être inférieur à 40 millisecondes). Donc cette méthode ne semble pas être la bonne pour détecter les polypes en temps réel. Cependant, on a pu voir que l'apprentissage machine utilisée ici permettait quand même d'avoir de bons résultats et il serait donc judicieux de l'utiliser par la suite.

C'est donc pour passer au dessus de cette limitation qu'une seconde méthode a été mise au point. Cette méthode basée uniquement sur de l'apprentissage machine pur devra permettre de détecter les polypes en temps réel. Pour effectuer l'apprentissage (qui donnera lieu à une classification), et comme pour n'importe quel type d'apprentissage, il faut absolument des images servant d'exemples, et c'est ce qui fait actuellement défaut pour les bases d'images de polypes. Afin de palier ce manque, on a mis au point un processus d'apprentissage actif qui permet d'améliorer nettement l'apprentissage machine. Plusieurs tests ont alors montré les bénéfices de cet apprentissage actif : le temps de calcul diminue et les performances de détection s'améliorent. On s'aperçoit ainsi, qu'avec les paramètres d'apprentissage basés sur les motifs binaires locaux, qu'on est capable d'effectuer la détection en temps réel quelque soit l'image après deux étapes d'apprentissage actif (on a un temps inférieur ou égal à 40 millisecondes). Avec de tels temps de calcul, la méthode est donc théoriquement capable de fonctionner en temps réel.

Avec une méthode capable de traiter des images en temps réel, on maintenant peut s'intéresser au traitement des vidéos. La détection dans les vidéos permet de se rapprocher de cas réels, et d'éprouver la méthode à un autre niveau. On a ainsi pu proposer une nouveau processus d'évaluation des performances en introduisant de nouvelles métriques qui sont cliniquement parlantes. Elles permettent indirectement de tenir compte du rendu à l'image de la détection. On a alors évalué notre méthode avec ces nouvelles métriques pour s'apercevoir que les performances de détection n'étaient pas si élevées qu'espérées. On a donc proposé plusieurs améliorations qui permettent d'améliorer les performances :

1. Fusion spatiale
2. Apprentissage actif
3. Cohérence temporelle
4. Retrait d'information
5. Retrait de la lumière spéculaire

Ces différentes améliorations tirent ainsi parti des optimisations mises en place pour

la détection des images fixes (apprentissage actif) et de l'utilisation des vidéos (cohérence temporelle). Mais on a aussi mis au point de nouvelles optimisations spécialement pour la détection dans les vidéos comme une fusion spatiale des zones détectées pour réduire les fausses détections, mais aussi des étapes de prétraitement sur l'image avant la détection (retrait des bordures noires sans information utile et retrait de la lumière spéculaire). Quelques soient les améliorations apportées, on a toujours été capable d'alerter le médecin sur la présence du polype dans chacune des vidéos, ce qui prouve bien l'intérêt d'une telle méthode pour la détection des polypes en coloscopie. Sur l'aspect capacité à effectuer les calculs en temps réel, suivant les optimisations, la méthode est capable de tourner en temps réel, ou s'approchant du temps réel (entre 23 et 42 millisecondes). Cependant, comme le montre l'annexe B, le temps de calcul est dépendant de la machine utilisée pour mener les tests, et en prenant une machine légèrement plus puissante, on sera capable d'effectuer la détection en temps réel avec toutes les optimisations.

On a donc pu mettre au point une méthode répondant théoriquement à tous les critères qu'on s'était imposés au début, c'est-à-dire une méthode permettant d'alerter le médecin sur la présence de polypes lors d'un examen coloscopique en temps réel. Concernant l'aspect embarquabilité, on a pris soin d'utiliser uniquement des fonctions qui répondaient à cet aspect. Il resterait donc maintenant à mettre au point un système intégrant la méthode pour créer un outil plug'n'play pour l'examen endoscopique.

## 4.2 Discussions

### 4.2.1 Données

L'un des plus gros problèmes rencontrés au cours de ces 4 dernières années est la quantité de données disponibles librement pour effectuer les étapes d'apprentissage et de test. En effet, ce faible nombre d'images peut remettre en question la précision des métriques utilisées, car elles dépendent directement du nombre d'images utilisées pour tester. Il faut donc créer une base de données conséquente d'images de polypes afin d'en améliorer la précision.

De plus, augmenter la quantité de données disponible permet aussi d'avoir un effet bénéfique sur la méthode : on a pu voir qu'on avait mis au point une stratégie d'apprentissage actif afin de palier le manque de données pour apprendre, et donc avec plus de données disponibles, on aura plus d'exemples d'apprentissage et on pourra améliorer sans aucun doute l'apprentissage de notre méthode pour la rendre plus performante, que ce soit en terme de performances de détection qu'en terme de performance en temps de calcul.

La mise au point d'une nouvelle base de données devra se faire d'un commun accord avec différents partenaires, car il existe déjà plusieurs bases de données, mais elles diffèrent toutes, que ce soit sur la résolution des images, l'appareil utilisé pour effectuer les captures, l'angle de prise de vue... Il faut donc un protocole de prise de vue d'images endoscopiques unique, représentatif de la réalité, afin d'obtenir des images de qualité et exploitables.

### 4.2.2 Portage de la méthode

Comme on peut le constater en annexe B, on s'est intéressé à l'implémentation matérielle de la méthode de détection des polypes sur différentes plateformes matérielles. Après plusieurs essais, on est arrivé à la conclusion que la Raspberry Pi3 semblait être la plateforme proposant le meilleur ratio temps de calcul - prix. Cependant, le temps de calcul proposé par la Raspberry Pi3 reste encore supérieur à la limite des 40 millisecondes à ne pas dépasser pour fonctionner en temps réel. On va donc proposer ici plusieurs idées afin de porter la méthode sur différentes plateformes et atteindre le temps réel.

#### Cluster de Raspberry Pi3

Comme on a pu le voir avec l'annexe B, la Raspberry Pi3 offre un très bon rapport qualité-prix et se prête bien à l'objectif de proposer un outil de détection des polypes adaptable à n'importe quel endoscope existant, offrant de bonnes performances, et dont le prix ne serait pas excessif afin de pouvoir l'embarquer dans la plupart des salles d'exams. Mais comme on a pu le constater, la méthode n'est pas capable de tourner en temps réel sur une carte Raspberry Pi3. Mais qu'en est-il si on en utilisait plusieurs au lieu d'une ?

C'est donc là qu'intervient le cluster formé à partir de plusieurs Raspberry Pi [CCB<sup>+</sup>14]. Un tel cluster permettrait de distribuer les images à traiter sur différentes Raspberry Pi3. En effet, si on dimensionne notre cluster pour traiter les images à un rythme de 25 images par seconde, on aura certes un décalage entre l'image en entrée du système et l'image en sortie du système, mais on sera capable de traiter au rythme de 25 images par secondes.

Pour dimensionner notre cluster, on va avoir besoin de plusieurs informations. Tout d'abord, il faut le temps moyen pour traiter une image sur Raspberry Pi3. Pour ce faire, on utilise la figure B.1 qui montre que le temps de traitement moyen d'une image est de 131 millisecondes. A ce moment là, on pourrait penser qu'il faudrait donc un cluster composé de 5 Raspberry Pi3 (une Raspberry Pi3 comme unité d'acquisition et d'affichage (UAA), et quatre Raspberry Pi3 comme unité de calculs (UC)). La justification d'une telle chose est simple : imaginons qu'à l'instant  $t$ , la carte 1 reçoit une image, elle va mettre 131 millisecondes à la traiter, il faut donc que la prochaine image arrive après 131 millisecondes. Donc au rythme de 40 millisecondes entre chaque image, le plus proche multiple de 40 supérieur à 131 est 160, soit  $4 \times 40$ . Donc il faudrait 4 Raspberry Pi3 comme unité de calculs. Cependant, cela ne marchera pas, car on n'a pas pris en compte les temps de transfert des images entre les unités de calculs et l'unité d'acquisition et d'affichage.

On a donc cherché à connaître ce temps de calcul afin de mieux dimensionner notre cluster. Pour y arriver, on effectue un test en anneau avec 4 Raspberry Pi3 (1 unité d'acquisition et d'affichage, et 3 unités de calculs). Un test en anneau consiste à envoyer depuis l'UAA un message vers l'UC1, puis ce message est envoyé de l'UC1 vers l'UC2, puis de l'UC2 vers l'UC3 et enfin de l'UC3 vers l'UAA. On mesure donc le temps que met le message pour revenir à l'UAA après son départ. En effectuant 10 essais, on arrive

à une vitesse de transfert des données moyenne dans le cluster de 10 Mo/s. Et comme les images de nos vidéos font en moyenne 320 ko, on peut donc en déduire qu'il faut environ 32 millisecondes pour transférer une image entre l'UAA et une UC (et vice versa). Ce qui veut dire qu'il faut 195 millisecondes ( $131 + 2 \times 32 = 195$ ) pour traiter une image sur une Raspberry Pi3 (temps de transferts compris). On aura donc une configuration du cluster comme sur la figure 4.1. Théoriquement, cela créera un cluster capable de traiter les images au rythme de 25 images par secondes, avec seulement un délai d'environ 200 millisecondes entre l'acquisition et l'affichage.

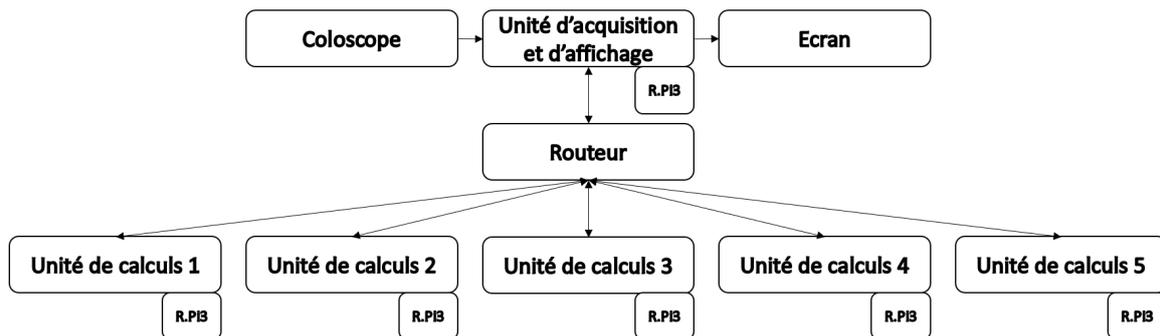


FIGURE 4.1 – Configuration du cluster pour la détection basée sur Raspberry Pi3 (R.Pi3)

### De l'interprété vers le compilé

Depuis le début, on a choisi d'utiliser le langage Python pour programmer la méthode de détection des polypes car il est pratique et flexible pour prototyper rapidement et portable sur différents systèmes d'exploitation. Ceci est rendu possible car il s'agit d'un langage interprété, et donc le code source est compatible avec n'importe quel interpréteur s'il est installé sur la machine utilisée. Cependant, le principal inconvénient est que, de manière générale, le langage interprété est plus lent à l'exécution que le langage compilé.

On pourrait donc s'intéresser à transformer le code de la méthode vers un langage compilé (comme du C ou du C++). En effet, au contraire d'un code interprété, le code compilé demandera à chaque fois une compilation sur la machine cible, mais sera plus rapide à l'exécution. On peut voir dans quelques exemples tirés de la littérature [Pre00] [FG08] que l'exécution du langage compilé est plus rapide et consomme moins de mémoire que le langage interprété.

De plus, certains langages peuvent tirer les avantages des deux types de langage (interprété et compilé) comme le Java ou le C# [FG08]. Il pourrait donc être intéressant de porter notre méthode dans l'un de ces deux langages afin d'améliorer nos temps de calcul et les rendre encore plus faibles.

## Du Python vers une implémentation matérielle sur FPGA

Pour pousser plus loin l'optimisation et l'embarquabilité de la méthode, on peut envisager de créer une IP (Intellectual Property) pour un FPGA (Field Programmable Gate Array) afin de créer un composant dédié à la détection des polypes et pouvant être, si besoin, intégré directement dans les systèmes endoscopiques (comme la capsule endoscopique ou les colonnes d'endoscopies).

Il semble très intéressant de faire une telle démarche car l'implémentation matérielle pourrait permettre d'accélérer le traitement des images. En effet, il est possible d'obtenir un temps de traitement par image plus petit que ce que l'on a pu observer jusqu'ici (jusqu'à plus de 100 images par seconde) pour la méthode de Viola et Jones quelque soit les paramètres d'apprentissage utilisés [LSC07] [CMOK09] [GL08] [HSM09] [WBC04].

Avec de tels outils, on serait capable de proposer un système miniature consommant moins d'énergie qu'un ordinateur et proposant des performances plus que convenables. C'est donc une piste à envisager pour la suite concernant l'intégration de l'outil.

### 4.2.3 Tracking des polypes pour réduire le temps de calcul et améliorer la détection

Afin d'améliorer les temps de calcul et alléger les plateformes utilisées pour effectuer les détections de polypes, et pour renforcer la cohérence temporelle, on peut envisager une solution de tracking qui permettrait d'éliminer un grand nombre de fausses détections.

La figure 4.2 montre le principe du tracking que l'on souhaite mettre en place, les zones en violet représentent les détections renvoyées par la méthode de détection des polypes, les zones en vert sont la vérité terrain et les zones en bleu représentent les zones suivies (tracking). Le but du tracking est alors de pouvoir quand même montrer la zone du polype alors que la détection n'a pas réussi à le trouver.

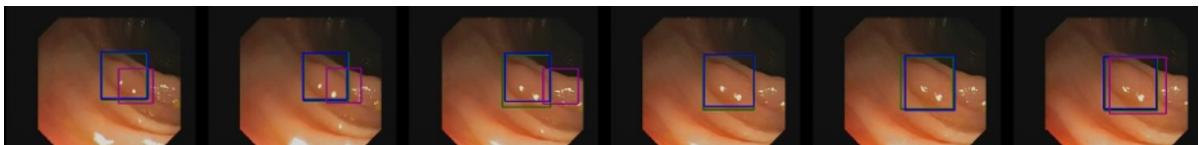


FIGURE 4.2 – Principe du tracking

Techniquement, ce tracking va être obtenu en faisant une extraction des caractéristiques temporelles, c'est-à-dire en se basant sur plusieurs images successives. Ces caractéristiques seront ainsi couplées à un apprentissage de type SVM (Support Vector Machine). Finalement, on utilisera la méthode de Camshift pour parvenir au tracking des polypes en supplément de la détection.

### 4.2.4 Essais cliniques

La prochaine étape pour la méthode est de participer à des essais cliniques. Ces essais permettraient enfin d'éprouver la méthode dans des conditions réelles. Grâce à ces essais, on pourra enfin voir notre méthode agir dans des conditions réelles, mais aussi, et le plus important, on pourra montrer l'applicabilité clinique. En effet, cela permettra d'avoir le retour des praticiens et d'identifier les points à améliorer. On pourra aussi effectuer des études croisées et voir en comparaison de l'approche du médecin, comment notre méthode permet d'améliorer réellement le taux de polypes manqués.

Pour réaliser ces essais, une collaboration avec l'Espagne a été mise en place. En particulier avec une équipe de chercheurs, dont Jorge Bernal, travaillant à l'UAB (Universitat Autònoma de Barcelona) sur la même thématique. Cela va ainsi permettre de réaliser des essais d'ici quelques mois à l'Hospital Clinic de Barcelone.

### 4.2.5 Vers la caractérisation des polypes colorectaux : une ouverture à la fluorescence

Pour aller encore plus loin dans le traitement des polypes, on pourrait envisager de les caractériser, c'est-à-dire savoir si le polype est bénin ou malin, et par conséquent savoir s'il est nécessaire de le retirer ou non. Pour effectuer cela, on s'intéresse à une méthode qui semble très prometteuse : la fluorescence.

#### Définition

Comme on a pu le voir avec la technologie AFI (pour AutoFluorescence Imaging), la fluorescence est une réaction de la matière avec la lumière. Lorsque la lumière réagit, on va observer une absorption de photons, qui va donner lieu à une transition électronique du niveau fondamental  $S_0$  vers les niveaux d'énergie supérieure  $S_1$ ,  $S_2$ ,  $S_n$ . Les niveaux électroniques supérieurs sont atteints si l'énergie absorbée est importante. Une reconversion énergétique interne permet de revenir au niveau  $S_1V_n$ . On passe alors du niveau  $V_n$  au niveau  $V_0$  grâce à l'interaction entre les molécules de fluorophore et les molécules de solvant. Ce n'est qu'à partir du niveau  $S_1V_0$  que peut se produire le phénomène de fluorescence comme indiqué sur la figure 4.3.

La fluorescence est l'émission de photons à partir d'un état singulet  $S_1$ . Les transitions entre l'état  $S_1$  vers l'état  $S_0$  se déroule avec des constantes de vitesse d'émission autour de  $10^{-8}$  secondes. L'intensité du spectre de la fluorescence dépend, en majeure partie, de la population de molécules à l'état excité  $S_1V_0$ . De plus, les processus qui entrent en compétition avec la fluorescence contribuent à faire diminuer cette population, et donc l'intensité du spectre lumineux.

Du point de vue énergétique, le chemin qui suit l'électron est fermé : il part de  $S_0V_0$  vers l'état  $S_nV_m$  (avec  $n$  et  $m$  strictement positifs), puis par reconversions internes, passe à l'état  $S_1V_0$ . Il passe ensuite à l'état  $S_0V_0$  avec émission de photons (la fluorescence). L'excitation apporte l'énergie  $E_{excitation} = \hbar\nu_{excitation}$  (avec  $\hbar$  la constante de Planck et

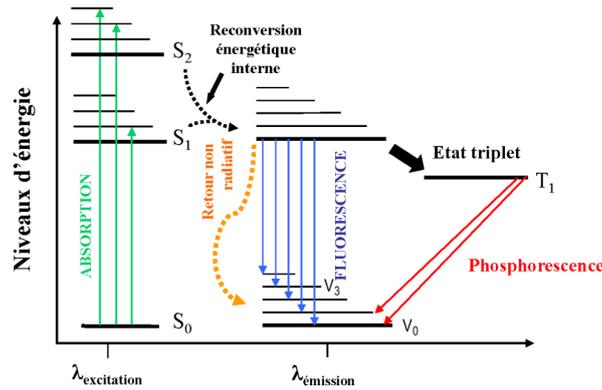


FIGURE 4.3 – Introduction à la fluorescence : diagramme de Jablonski

$\nu_{excitation}$  la fréquence d'excitation). L'électron va donc restituer cet apport d'énergie par les différentes reconversions internes et par l'énergie associée au photon de fluorescence. De plus, comme le montre l'équation 4.1, la longueur d'onde émise par la fluorescence sera forcément supérieure à la longueur d'onde utilisée pour l'excitation.

$$E_{excitation} \geq E_{émission} \Rightarrow \nu_{excitation} \geq \nu_{émission} \Rightarrow \lambda_{émission} \geq \lambda_{excitation} \quad (4.1)$$

Cependant, cette fluorescence n'est pas permanente, elle possède un temps de vie. Il est égal au temps moyen pendant lequel les molécules restent à l'état excité avant de revenir à l'état fondamental. L'ordre de grandeur de ce temps de vie varie de la nanoseconde à la picoseconde. Une fluorophore peut avoir un ou plusieurs temps de vie de fluorescence (qui est le cas pour les protéines par exemple).

Dans le domaine médical, il existe principalement deux types de fluorescence : la fluorescence exogène et la fluorescence endogène. L'application à ce domaine de la fluorescence permet de caractériser les lésions détectées par le médecin :

- Les lésions hyperplasiques qui ne sont pas dangereuses
- Les lésions préneoplasiques qui peuvent donner lieu à un cancer
- Les adénomes qui sont des cellules cancéreuses

Certains fluorophores permettent donc de mettre en évidence ces différentes lésions.

### La fluorescence exogène

Ce type de fluorescence est obtenu à partir de fluorophores exogènes, c'est-à-dire qui ne sont pas présents dans l'organisme du patient. ces fluorophores sont principalement utilisés en tant qu'agent de contraste pour la détection de cancer incluant des photosensibilisants, développés à la base pour la thérapie photodynamique. L'avantage de ce genre de fluorophores est de connaître la longueur d'onde pour les exciter mais aussi la longueur

| Fluorophore                | Excitation (nm) | Emission (nm)     |
|----------------------------|-----------------|-------------------|
| <i>HP</i>                  | 405             | 610               |
| <i>BpD – MA</i>            | 400             | 690               |
| <i>HpD</i>                 | 400             | 610               |
| <i>MACE</i>                | 410             | 664               |
| <i>Sn.NT2H<sub>2</sub></i> | 410             | 645               |
| <i>AISPC</i>               | 350             | 675               |
| <i>NBA</i>                 | 620             | 660               |
| <i>Indocyanine</i>         | 738             | 690-720 & 760-800 |

TABLE 4.1 – Liste des fluorophores exogènes

d'onde à laquelle opère la fluorescence.

La tableau 4.1 montre les différents types de fluorophores avec les longueurs d'onde d'excitation et d'émission de la fluorescence. [AEW92] [FWT03]

On constate ici que la plupart des agents de contrastes utilisés ont une longueur d'onde de fluorescence qui se situe dans le rouge.

On note aussi l'expérimentation du lutétium texaphyrine, un agent photosensibilisant qui présente des caractéristiques d'absorption et de fluorescence dans le proche infrarouge (NIR, Near InfrRed). Cet agent semble se concentrer dans les tumeurs grâce à l'apoptose plutôt que la nécrose, donc dans la cellule même. Ce qui permettrait une meilleure localisation des tumeurs.

Cependant, l'injection d'agents de contraste à un patient reste un risque car ces substances ne sont pas naturellement présentes dans son organisme. De plus, il faut un certain temps au fluorophore pour rejoindre les cellules afin d'obtenir le meilleur contraste. Suivant le fluorophore, cela peut varier entre 2h et 24h. [AEW92]

### La fluorescence endogène

La fluorescence endogène (ou autofluorescence) est, contrairement à la fluorescence exogène, naturelle car les fluorophores sont produits par le corps humain. Il n'y a donc pas besoin de faire une injection d'un fluorophore au patient.

Le tableau 4.2 permet de résumer les différents types de fluorophores endogènes existants ainsi que les longueurs d'ondes d'excitation et d'émission. [AEW92] [Ram00] [BDV<sup>+</sup>06] [RBG<sup>+</sup>13]

L'excitation en proche infrarouge peut aussi donner lieu à de l'autofluorescence. Une excitation avec un laser d'une longueur d'onde de 785 nm provoque une autofluorescence des tissus dans la bande 810-1050nm et l'intensité de la fluorescence est liée à la condition

| Fluorophore             | Excitation (nm)       | Emission (nm)         |
|-------------------------|-----------------------|-----------------------|
| <i>Tryptophan</i>       | 280                   | 350                   |
| <i>Tyrosine</i>         | 275                   | 300                   |
| <i>Phenylalanine</i>    | 260                   | 280                   |
| <i>Collagen</i>         | 340 & 270 & 285       | 395 & 395 & 310       |
| <i>Elastin</i>          | 460 & 360 & 425 & 260 | 520 & 410 & 490 & 410 |
| <i>NADH</i>             | 290 & 350             | 440 & 460             |
| <i>NADPH</i>            | 336                   | 464                   |
| <i>FAD</i>              | 450                   | 535                   |
| <i>FMN</i>              | 440                   | 520                   |
| <i>VitamineA</i>        | 327                   | 510                   |
| <i>VitamineK</i>        | 335                   | 480                   |
| <i>VitamineD</i>        | 390                   | 480                   |
| <i>Pyridoxine</i>       | 332 & 340             | 400                   |
| <i>Pyridoxamine</i>     | 335                   | 400                   |
| <i>Pyridoxal</i>        | 330                   | 385                   |
| <i>Acidepyridoxique</i> | 315 & 330             | 425 & 400             |
| <i>VitamineB12</i>      | 275                   | 305                   |
| <i>Phospholipides</i>   | 436                   | 540 & 560             |
| <i>Lipofuscine</i>      | 340-395               | 540 & 430-460         |
| <i>Céroïde</i>          | 340-395               | 430-460 & 540         |
| <i>Porphyrynes</i>      | 400                   | 610 & 675             |

TABLE 4.2 – Liste des fluorophores endogènes

du tissu (cancéreux ou non) [SZH11]. En effet, les tissus montrent un pic d'intensité lumineuse vers 810 nm mais plus le tissu est infecté, plus cette intensité diminue. L'origine de cette baisse d'intensité est liée aux porphyrines qui possèdent une caractéristique d'absorption dans le proche infrarouge importante et fluorescent dans l'infrarouge éloigné [ZDA99].

L'utilisation de l'excitation infrarouge (IR) est aussi abordée dans différents articles. On note principalement que l'excitation dans les bandes de longueurs d'ondes suivantes, 1226-1251 nm et 1288-1370 nm [AUK<sup>+</sup>11], permet d'obtenir le meilleur contraste entre cellules saines et cellules infectées, cependant, il s'agit d'analyse hyperspectrale. Un deuxième constat [NLT14] permet de mettre en évidence un triplet de longueurs d'onde qui permettent d'obtenir un contraste idéal : 1100 nm, 1220 nm et 1580 nm. L'excitation en infrarouge permet donc aussi de caractériser les tissus.

### Liste des abréviations

- *HP* : Hématoporphyrine
- *BpD – MA* : Benzoporphyrine monoacétate
- *HpD* : Dérivé d'hématoporphyrine
- *MACE (NPe6)* : Mono-L-aspartyle chlorure e6
- *Sn.NT2H<sub>2</sub>* : Etain réduit métallo purpurine
- *AISPC* : Tétra sulfoné de phtalocyanine d'aluminium
- *NBA* : Bleu de Nel A
- *NADH* : Nicotinamide adénine dinucléotide
- *NADPH* : Nicotinamide adénine dinucléotide phosphate
- *FAD* : Flavine adénine dinucléotide
- *FMN* : Flavin mononucléotide

### Application clinique

Actuellement, l'autofluorescence est utilisée en médecine pour la détection et la caractérisation des lésions : il s'agit du système AFI (AutoFluorescence Imaging). Son principe de fonctionnement est le suivant : il excite dans le domaine des ultraviolets les différents tissus à analyser puis, avec une caméra, regarde la fluorescence de ces tissus (avec la mesure de certaines composantes de couleur). Des méthodes automatiques pour détecter les polypes y sont associées comme par exemple une analyse de couleurs avec seuil à franchir pour caractériser le polype comme dangereux [ITA<sup>+</sup>13]. L'autofluorescence présente un gros avantage par rapport à la coloscopie en lumière blanche car elle permet d'augmenter le taux de détection des adénomes [MFS<sup>+</sup>12] pour les gastroentérologues débutants. [MSF<sup>+</sup>08] constate que le taux de lésions manquées lors d'un examen diminue avec l'AFI comparé à la coloscopie en lumière blanche mais souligne le fait que la préparation du patient joue un rôle essentiel.

Cela montre donc que l'utilisation de l'autofluorescence pourrait être un bon complément à notre outil de détection des polypes pour apporter une caractérisation des polypes de manière non invasive. Il suffirait alors de trouver la longueur d'onde adaptée pour la

caractérisation des polypes.

# Liste des publications

## 1. Revue

- *Comparative Validation of Polyp Detection Methods in Video Colonoscopy : Results from the MICCAI 2015 Endoscopic Vision Challenge* - Jorge Bernal, Nima Tajkbaksh, F Sánchez, Bogdan Matuszewski, Hao Chen, Lequan Yu, **Quentin Angermann**, Olivier Romain, Bjørn Rustad, Ilangko Balasingham, Konstantin Pogorelov, Sungbin Choi, Quentin Debard, Lena Maier-Hein, Stefanie Speidel, Danail Stoyanov, Patrick Brandao, Henry Córdova, Cristina Sánchez-Montes, Suryakanth Gurudu, Gloria Fernández-Esparrach, Xavier Dray, Jianming Liang and Aymeric Histace - 2017 - IEEE Transactions on Medical Imaging, Institute of Electrical and Electronics Engineers - 36 (6) - pp.1231 - 1249

## 2. Brevet

- *European Patent : Application Method and Apparatus for Real-Time Detection of Polyps in Optical Colonoscopy* - **Quentin Angermann**, Aymeric Histace, Olivier Romain and Xavier Dray - 2018 - Patent n° EP3267394A1

## 3. Conférences internationales

- *Towards Real-Time Polyp Detection in Colonoscopy Videos : Adapting Still Frame-Based Methodologies for Video Sequences Analysis* - **Quentin Angermann**, Jorge Bernal, Cristina Sánchez - Montes, Maroua Hammami, Gloria Fernández - Esparrach, Xavier Dray, Olivier Romain, Javier Sanchez and Aymeric Histace - 2017 - Springer - CARE 2017 (in conjunction with MICCAI 2017) - Québec City, Canada - Lecture Notes on Computer Science, LNCS 10550 - pp.29-41 - Proceedings of Computer Assisted and Robotic Endoscopy Workshop
- *Clinical Usability Quantification Of a Real-Time Polyp Detection Method In Videocolonoscopy* - **Quentin Angermann**, Jorge Bernal, Cristina Sánchez-Montes, Maroua Hammami, Gloria Fernández-Esparrach, Xavier Dray, Olivier Romain, Javier Sanchez and Aymeric Histace - 2017 - 25th United European Gastroenterology Week - Barcelona, Spain - Proceedings of UEGW conference
- *Hardware Platforms Benchmark For Real-Time Polyp Detection* - **Quentin Angermann**, Aymeric Histace, Maroua Hammami, Mehdi Terosiet, Lionel

Faurlini and Olivier Romain - 2017 - Euromicro Conference on Digital System Design - Vienna, Austria - Proceedings of DSD conference

- *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, Javier Sanchez and Aymeric Histace - 2017 - International Journal of Computer-Assisted Radiology and Surgery - Barcelona, Spain - Springer, Proceedings of CARS conference
- *Embeddable Real Time Tool For Automatic Skin Lesions Characterization* - **Quentin Angermann**, Maroua Hammami, Aymeric Histace and Olivier Romain - 2017 - Proceedings of COMET 2017 Conference - Cergy, France
- *Apprentissage actif pour la détection en temps réel des polypes en coloscopie* - **Quentin Angermann**, Aymeric Histace et Olivier Romain - 2016 - 1er Forum Franco - Québécois d'Innovation en Santé - Montréal, Canada
- *Active Learning For Real Time Detection Of Polyps In Videocolonoscopy* - **Quentin Angermann**, Aymeric Histace and Olivier Romain - Elsevier, Medical Image Understanding and Analysis Conference - Loughborough, United Kingdom - 90 - pp.182-187 - Procedia Computer Science
- *Real-Time Tracking Of Polyps For Computer-Assisted Colonoscopy* - **Quentin Angermann**, Aymeric Histace, Olivier Romain and Xavier Dray - 2015 - International Conference on Computer-Assisted Radiology and Surgery - Barcelona, Spain - Proceedings of 29th CARS conference - International Journal of CARS Supp
- *Cyclope : A Smart Videoapsule For Early Diagnosis of Colorectal Cancer* - **Quentin Angermann**, Aymeric Histace, Olivier Romain, Bertrand Granado and Xavier Dray - 2014 - International Symposium on Biomaterial and Smart Systems

#### 4. Chapitre d'ouvrage

- *Smart Videocapsule for Early Diagnosis of Colorectal Cancer : Toward Embedded Image Analysis* - **Quentin Angermann**, Aymeric Histace, Olivier Romain, Xavier Dray, Andrea Pinna and Bertrand Granado - 2015 - Springer - Computational Intelligence in Digital and Network Designs and Applications - 2 (Chapter 12) - pp.325-350 - Part 2 : Digital, Network Designs and Applications

#### 5. Journées GdR

- *Real-Time Tracking Of Polyps For Computer-Assisted Colonoscopy* - **Quentin Angermann**, Aymeric Histace, Olivier Romain and Xavier Dray - 2015 - GdR

ISIS (Information, Signal, Image et ViSion)



# Bibliographie

- [ABSM<sup>+</sup>17a] Quentin Angermann, Jorge J Bernal, Cristina Sánchez-Montes, Maroua Hammami, Gloria Fernández-Esparrach, Xavier Dray, Olivier Romain, Javier Sanchez, and Aymeric Histace. Clinically Usability Quantification Of a Real-Time Polyp Detection Method In Videocolonoscopy. In *United European Gastroenterology Journal*, Proceedings of 25th UEG week, 2017.
- [ABSM<sup>+</sup>17b] Quentin Angermann, Jorge J Bernal, Cristina Sánchez-Montes, Maroua Hammami, Gloria Fernández-Esparrach, Xavier Dray, Olivier Romain, Javier Sanchez, and Aymeric Histace. Real-Time Polyp Detection in Colonoscopy Videos : A Preliminary Study For Adapting Still Frame-based Methodology To Video Sequences Analysis. In *International Journal of Computer-Assisted Radiology and Surgery*, Proceedings of CARS conference, 2017.
- [ADW<sup>+</sup>13] Wendy Atkin, Edward Dadswell, Kate Wooldrage, Ines Kralj-Hans, Christian von Wagner, Rob Edwards, Guiqing Yao, Clive Kay, David Burling, Omar Faiz, et al. Computed tomographic colonography versus colonoscopy for investigation of patients with symptoms suggestive of colorectal cancer (siggar) : a multicentre randomised trial. *The Lancet*, 381(9873) :1194 – 1202, 2013.
- [AEW92] Stefan Andersson-Engels and Brian C Wilson. In vivo fluorescence in clinical oncology : fundamental and practical issues. *J Cell Pharmacol*, 3(1) :48, 1992.
- [AHR17] Quentin Angermann, Maroua Hammami, Aymeric Histace, and Olivier Romain. Embeddable Real Time Tool For Automatic Skin Lesions Characterization. In *Proceedings of COMET 2017 Conference*, 2017.
- [AHR<sup>+</sup>15] Quentin Angermann, Aymeric Histace, Olivier Romain, Xavier Dray, Andrea Pinna, and Bertrand Granado. Smart videocapsule for early diagnosis of colorectal cancer : toward embedded image analysis. In *Computational Intelligence in Digital and Network Designs and Applications*, pages 325 – 350. 2015.
- [AHR16] Quentin Angermann, Aymeric Histace, and Olivier Romain. Active learning for real time detection of polyps in videocolonoscopy. *Procedia Computer Science*, 90 :182 – 187, 2016.
- [Alb01] Jihad René Albani. Absorption et fluorescence. 2001.
- [AUK<sup>+</sup>11] Hamed Akbari, Kuniaki Uto, Yukio Kosugi, Kazuyuki Kojima, and Naofumi Tanaka. Cancer detection using infrared hyperspectral imaging. *Cancer science*, 102(4) :852–857, 2011.

- [AWP<sup>+</sup>09] Stefan Ameling, Stephan Wirth, Dietrich Paulus, Gerard Lacey, and Fernando Vilarino. Texture-based polyp detection in colonoscopy. *Bildverarbeitung für die Medizin 2009*, pages 346 – 350, 2009.
- [BB10] Steven R Brown and Wal Baraza. Chromoscopy versus conventional endoscopy for the detection of polyps in the colon and rectum. *Cochrane Database Syst Rev*, 10, 2010.
- [BDV<sup>+</sup>06] WCPM Blondel, Ch Daul, Sandrine Villette, R Miranda-Luna, Yahir Hernandez-Mier, Geneviève Bourg-Heckly, and Didier Wolf. Spectro-imagerie endoscopique d'autofluorescence et cartographie 2d pour la localisation et le diagnostic in situ de lésions cancéreuses. 2006.
- [BfCCS<sup>+</sup>10] D Burling, International Collaboration for CT Colonography Standards, et al. Ct colonography standards. *Clinical radiology*, 65(6) :474 – 480, 2010.
- [BHM<sup>+</sup>18] Jorge Bernal, Aymeric Histace, Marc Masana, Quentin Angermann, Cristina Sánchez-Montes, Cristina Rodriguez de Miguel, Maroua Hammami, Ana Garcia-Rodriguez, Henry Códova, Olivier Romain, Gloria Fernández-Esparrach, Xavier Dray, and Javier Sanchez. Polyp detection benchmark in colonoscopy videos using gtcreator : A novel fully configurable tool for easy and fast annotation of image databases. In *International Congress and Exhibition on Computer Assisted Radiology and Surgery*, Proceedings of CARS conference, 2018.
- [BSFE<sup>+</sup>15] Jorge Bernal, F Javier Sánchez, Gloria Fernández-Esparrach, Debora Gil, Cristina Rodríguez, and Fernando Vilariño. Wm-dova maps for accurate polyp highlighting in colonoscopy : Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics*, 43 :99 – 111, 2015.
- [BSV12] Jorge Bernal, Javier Sánchez, and Fernando Vilarino. Towards automatic polyp detection with a polyp appearance model. *Pattern Recognition*, 45(9) :3166 – 3182, 2012.
- [BSV13] Jorge Bernal, Javier Sánchez, and Fernando Vilarino. Impact of image preprocessing methods on polyp localization in colonoscopy frames. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 7350–7354, 2013.
- [BY15] Seung-Hwan Bae and Kuk-Jin Yoon. Polyp detection via imbalanced learning and discriminative feature learning. *IEEE transactions on medical imaging*, 34(11) :2379 – 2393, 2015.
- [CCB<sup>+</sup>14] Simon J Cox, James T Cox, Richard P Boardman, Steven J Johnston, Mark Scott, and Neil S O'brien. Iridis-pi : a low-cost, compact demonstration cluster. *Cluster Computing*, 17(2) :349–358, 2014.
- [CMOK09] Junguk Cho, Shahnam Mirzaei, Jason Oberg, and Ryan Kastner. Fpga-based face detection system using haar classifiers. In *Proceedings of the ACM/SIGDA International Symposium on Field Programmable Gate Arrays*, FPGA '09, pages 103–112, New York, NY, USA, 2009. ACM.
- [CP11] Joseph Coombs and R Prabhu. Opencv on ti's dsp+ arm® platforms : Mitigating the challenges of porting opencv to embedded platforms. *Texas Instruments*, 2011.

- [Dav88] ER Davies. A modified hough scheme for general circle location. *Pattern Recognition Letters*, 7(1) :37–43, 1988.
- [Der10] Konstantinos G Derpanis. Overview of the ransac algorithm. *Image Rochester NY*, 4(1) :2 – 3, 2010.
- [DH72] Richard O Duda and Peter E Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1) :11–15, 1972.
- [DMS14] Vincent K Dik, Leon MG Moons, and Peter D Siersema. Endoscopic innovations to increase the adenoma detection rate during colonoscopy. *World Journal of Gastroenterology*, 20(9) :2200 – 2211, 2014.
- [FG08] Mathieu Fourment and Michael R Gillings. A comparison of common programming languages used in bioinformatics. *BMC bioinformatics*, 9(1) :82, 2008.
- [Fre95] Yoav Freund. Boosting a weak learning algorithm by majority. *Information and computation*, 121(2) :256 – 285, 1995.
- [FSE<sup>+</sup>15] J Ferlay, I Soerjomataram, M Ervik, R Dikshit, S Eser, C Mathers, M Rebelo, DM Parkin, D Forman, and F Bray. Cancer incidence and mortality worldwide : Iarc cancerbase no. 11 [internet]. lyon, france : International agency for research on cancer. globocan. 2013 ; 2012 v1. 0. Available from : <http://globocan.iarc.fr>, 2015.
- [FTC<sup>+</sup>07] Joshua J. Fenton, Stephen H. Taplin, Patricia A. Carney, Linn Abraham, Edward A. Sickles, Carl D’Orsi, Eric A. Berns, Gary Cutter, R. Edward Hendrick, William E. Barlow, and Joann G. Elmore. Influence of computer-aided detection on performance of screening mammography. *New England Journal of Medicine*, 356(14) :1399–1409, 2007.
- [Fuj07] Fujinon Corporation. *FICE Catalog*, 2007.
- [FWT03] Martin Funovics, Ralph Weissleder, and Ching-Hsuan Tung. Protease sensors for bioimaging. *Analytical and bioanalytical chemistry*, 377(6) :956–963, 2003.
- [GBC<sup>+</sup>11] SA Gross, AM Buchner, JE Crook, JR Cangemi, Michael F Picco, Herbert C Wolfsen, KR DeVault, DS Loeb, Massimo Raimondo, Timothy A Woodward, et al. A comparison of high definition-image enhanced colonoscopy and standard white-light colonoscopy for colorectal polyp detection. *Endoscopy*, 43(12) :1045–1051, 2011.
- [GCC<sup>+</sup>16] David Gutman, Noel CF Codella, Emre Celebi, Brian Helba, Michael Marchetti, Nabin Mishra, and Allan Halpern. Skin lesion analysis toward melanoma detection : A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv :1605.01397*, 2016.
- [GL08] Changjian Gao and Shih-Lien Lu. Novel fpga based haar classifier face detection algorithm acceleration. In *Field Programmable Logic and Applications, 2008. FPL 2008. International Conference on*, pages 373–378. IEEE, 2008.

- [GSB<sup>+</sup>09] Sebastian Gross, Thomas Stehle, Alexander Behrens, Roland Auer, Til Aach, Ron Winograd, Christian Trautwein, and Jens Tischendorf. A comparison of blood vessel features and local binary patterns for colorectal polyp classification. *Proc. of Medical Imaging*, 2009.
- [GSH<sup>+</sup>14] Ian M Gralnek, Peter D Siersema, Zamir Halpern, Ori Segol, Alaa Melhem, Alain Suissa, Erwin Santo, Alan Sloyer, Jay Fenster, Leon MG Moons, et al. Standard forward-viewing colonoscopy versus full-spectrum endoscopy : an international, multicentre, randomised, tandem colonoscopy trial. *The lancet oncology*, 15(3) :353 – 360, 2014.
- [HOT<sup>+</sup>07] Sae Hwang, JungHwan Oh, Wallapak Tavanapong, Johnny Wong, and Piet C De Groen. Polyp detection in colonoscopy video using elliptical shape feature. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 2, page 465, 2007.
- [Hou62] Paul VC Hough. Method and means for recognizing complex patterns. Technical report, 1962.
- [HS<sup>+</sup>73] Robert M Haralick, Karthikeyan Shanmugam, et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6) :610 – 621, 1973.
- [HSM09] Masayuki Hiromoto, Hiroki Sugano, and Ryusuke Miyamoto. Partially parallel architecture for adaboost-based detection with haar-like features. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(1) :41–52, 2009.
- [Hu62] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 8(2) :179 – 187, 1962.
- [ISH<sup>+</sup>13] Yuji Iwahori, Takayuki Shinohara, Akira Hattori, Robert J Woodham, Shinji Fukui, Manas Kamal Bhuyan, and Kunio Kasugai. Automatic polyp detection in endoscope images using a hessian filter. In *MVA*, pages 21 – 24, 2013.
- [ITA<sup>+</sup>13] Hiroko Inomata, Naoto Tamai, Hiroyuki Aihara, Kazuki Sumiyama, Shoichi Saito, Tomohiro Kato, and Hisao Tajiri. Efficacy of a novel auto-fluorescence imaging system with computer-assisted color analysis for assessment of colorectal lesions. *World journal of gastroenterology : WJG*, 19(41) :7146, 2013.
- [KB09] Alexandros Karargyris and Nikolaos Bourbakis. Identification of polyps in wireless capsule endoscopy videos using log gabor filters. In *Life Science Systems and Applications Workshop, 2009. LiSSA 2009. IEEE/NIH*, pages 143–147. IEEE, 2009.
- [KBV<sup>+</sup>04] Ralf Kiesslich, Juergen Burg, Michael Vieth, Janina Gnaendiger, Meike Enders, Peter Delaney, Adrian Polglase, Wendy McLaren, Daniela Janell, Steven Thomas, et al. Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo. *Gastroenterology*, 127(3) :706 – 713, 2004.
- [KD03] J Kang and R Doraiswami. Real-time image processing system for endoscopic applications. In *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, volume 3, pages 1469 – 1472, 2003.

- [KIM<sup>+</sup>03] Stavros A Karkanis, Dimitrios K Iakovidis, Dimitrios E Maroulis, Dimitris A. Karras, and M Tzivras. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE transactions on information technology in biomedicine*, 7(3) :141 – 152, 2003.
- [Koh95] Teuvo Kohonen. Learning vector quantization. In *Self-Organizing Maps*, pages 175–189. Springer, 1995.
- [Kov00] Peter Kovesi. What are log-gabor filters and why are they good? *School of Computer Science & Software Engineering, The University of Western Australia*, 2000.
- [KTN<sup>+</sup>96] Shin-Ei Kudo, Satoru Tamura, Takashi Nakajima, Hiro-o Yamano, Hisashi Kusaka, and Hidenobu Watanabe. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointestinal endoscopy*, 44(1) :8–14, 1996.
- [LHN<sup>+</sup>06] M-G Lapalus, T Helbert, B Napoleon, JF Rey, P Houcke, and T Ponchon. Does chromoendoscopy with structure enhancement improve the colonoscopic adenoma detection rate? *Endoscopy*, 38(05) :444 – 448, 2006.
- [LLLM86] Hungwen Li, Mark A Lavin, and Ronald J Le Master. Fast hough transform : A hierarchical approach. *Computer Vision, Graphics, and Image Processing*, 36(2-3) :139–161, 1986.
- [LM02] Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 1, 2002.
- [LM12] Baopu Li and Max Q-H Meng. Automatic polyp detection for wireless capsule endoscopy images. *Expert Systems with Applications*, 39(12) :10952 – 10958, 2012.
- [LSC07] H. C. Lai, M. Savvides, and T. Chen. Proposed fpga hardware architecture for high frame rate (»100 fps) face detection using feature cascade classifiers. In *2007 First IEEE International Conference on Biometrics : Theory, Applications, and Systems*, pages 1–6, 2007.
- [LYS<sup>+</sup>13] Lee Guan Lim, Khay Guan Yeoh, Supriya Srivastava, Yiong Huak Chan, Ming Teh, and Khek Yu Ho. Comparison of probe-based confocal endomicroscopy with virtual chromoendoscopy and white-light endoscopy for diagnosis of gastric intestinal metaplasia. *Surgical endoscopy*, 27(12) :4649–4655, 2013.
- [LZL<sup>+</sup>07] Shengcai Liao, Xiangxin Zhu, Zhen Lei, Lun Zhang, and Stan Li. Learning multi-scale block local binary patterns for face recognition. *Advances in Biometrics*, pages 828 – 837, 2007.
- [MAAJ13] Ammara Masood and Adel Ali Al-Jumaily. Computer aided diagnostic support system for skin cancer : a review of techniques and algorithms. *International journal of biomedical imaging*, 2013, 2013.
- [MCE<sup>+</sup>16] Michael A Manfredi, Shailendra S Chauhan, Brintha K Enestvedt, Larissa L Fujii-Lau, Vani Konda, Joo Ha Hwang, Simon K Lo, John T Maple, Faris M Murad, Rahul Pannala, et al. Endoscopic electronic medical record systems. *Gastrointestinal endoscopy*, 83(1) :29–36, 2016.

- [MFFT14] Alexander V Mamonov, Isabel N Figueiredo, Pedro N Figueiredo, and Yen-Hsi Richard Tsai. Automated polyp detection in colon capsule endoscopy. *IEEE transactions on medical imaging*, 33(7) :1488 – 1502, 2014.
- [MFS<sup>+</sup>12] Kentaro Moriichi, Mikihiro Fujiya, Ryu Sato, Jiro Watari, Yoshiki Nomura, Toshie Nata, Nobuhiro Ueno, Shigeaki Maeda, Shin Kashima, Kentaro Itabashi, et al. Back-to-back comparison of auto-fluorescence imaging (afi) versus high resolution white light colonoscopy for adenoma detection. *BMC gastroenterology*, 12(1) :75, 2012.
- [MGH<sup>+</sup>16] Firmino M, Angelo G, Morais H, Dantas MR, and Valentim R. Computer-aided detection (cade) and diagnosis (cadx) system for lung cancer with likelihood of malignancy. *BioMedical Engineering OnLine*, 15(1), 2016.
- [MJR<sup>+</sup>13] Gaurav Mitra, Beau Johnston, Alistair P Rendell, Eric McCreath, and Jun Zhou. Use of simd vector operations to accelerate application code performance on low-powered arm and intel platforms. In *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2013 IEEE 27th International*, pages 1107–1116. IEEE, 2013.
- [MSF<sup>+</sup>08] Takahisa Matsuda, Yutaka Saito, Kuang-I Fu, Toshio Uraoka, Nozomu Kobayashi, Takeshi Nakajima, Hisatomo Ikehara, Yumi Mashimo, Tadakazu Shimoda, Yoshitaka Murakami, et al. Does autofluorescence imaging videoendoscopy system improve the colonoscopic polyp detection rate? - a pilot study. *The American journal of gastroenterology*, 103(8) :1926, 2008.
- [NLT14] Dorra Nouri, Yves Lucas, and Sylvie Treuillet. Infrared hyperspectral imaging for enhanced tissue visualization and discrimination during surgical operation. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 5137–5141. IEEE, 2014.
- [OPH96] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1) :51 – 59, 1996.
- [Pre00] Lutz Prechelt. An empirical comparison of c, c++, java, perl, python, rexx and tcl. *IEEE Computer*, 33(10) :23–29, 2000.
- [PRW11] Abraham M. Panossian, Massimo Raimondo, and Herbert C. Wolfsen. State of the art in the endoscopic imaging and ablation of barrett’s esophagus. *Digestive and Liver Disease*, 43(5) :365 – 373, 2011.
- [PS16] Sun Young Park and Dusty Sargent. Colonoscopic polyp detection using convolutional neural networks. In *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016.
- [Ram00] Nirmala Ramanujam. Fluorescence spectroscopy of neoplastic and non-neoplastic tissues. *Neoplasia*, 2(1) :89–117, 2000.
- [RBG<sup>+</sup>13] Timothy E Renkoski, Bhaskar Banerjee, Logan R Graves, Nathaniel S Rial, Sirandon AH Reid, Vassiliki L Tsikitis, Valentine N Nfonsam, Piyush Tiwari, Hemanth Gavini, and Urs Utzinger. Ratio images and ultraviolet c excitation in autofluorescence imaging of neoplasms of the human colon. *Journal of Biomedical Optics*, 18(1) :016005, 2013.

- [RHP<sup>+</sup>10] Dewkoemar Ramsoekh, Jelle Haringsma, Jan Werner Poley, Paul van Putten, Herman van Dekken, Ewout W Steyerberg, Monique E van Leerdam, and Ernst J Kuipers. A back-to-back comparison of white light video endoscopy with autofluorescence endoscopy for adenoma detection in high-risk subjects. *Gut*, 59(6) :785 – 793, 2010.
- [RKH14] Johannes Wilhelm Rey, Ralf Kiesslich, and Arthur Hoffman. New aspects of modern endoscopy. *World journal of gastrointestinal endoscopy*, 6(8) :334, 2014.
- [RUH16] Eduardo Ribeiro, Andreas Uhl, and Michael Häfner. Colonic polyp classification with convolutional neural networks. In *Computer-Based Medical Systems (CBMS), 2016 IEEE 29th International Symposium on*, pages 253 – 258, 2016.
- [ŠAB<sup>+</sup>16] I Ševo, A Avramović, Ilangko Balasingham, Ole Jakob Elle, Jacob Bergsland, and Lars Aabakken. Edge density based automatic detection of inflammation in colonoscopy videos. *Computers in biology and medicine*, 72 :138 – 150, 2016.
- [SB97] Stephen M Smith and J Michael Brady. Susan, a new approach to low level image processing. *International journal of computer vision*, 23(1) :45–78, 1997.
- [Sch90] Robert E. Schapire. The strength of weak learnability. *Mach. Learn.*, 5(2) :197 – 227, 1990.
- [SHMN<sup>+</sup>11] Cristiano Spada, Cesare Hassan, Miguel Munoz-Navas, Horst Neuhaus, Jacques Deviere, Paul Fockens, Emmanuel Coron, Gerard Gay, Ervin Toth, Maria Elena Riccioni, et al. Second-generation colon capsule endoscopy compared with colonoscopy. *Gastrointestinal endoscopy*, 74(3) :581 – 589, 2011.
- [SHR<sup>+</sup>14] Juan S Silva, Aymeric Histace, Olivier Romain, Xavier Dray, and Bertrand Granado. Towards embedded detection of polyps in wce images for early diagnosis of colorectal cancer. *International Journal of Computer Assisted Radiology and Surgery*, 9(2) :283 – 293, 2014.
- [Soc15] American Cancer Society. Global cancer facts and figures, 3rd edition, 2015. Atlanta : American Cancer Society.
- [SZH11] Xiaozhuo Shao, Wei Zheng, and Zhiwei Huang. Near-infrared autofluorescence spectroscopy for in vivo identification of hyperplastic and adenomatous polyps in the colon. *Biosensors and Bioelectronics*, 30(1) :118–122, 2011.
- [TGL16] Nima Tajbakhsh, Suryakanth R Gurudu, and Jianming Liang. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE transactions on medical imaging*, 35(2) :630 – 644, 2016.
- [TYK<sup>+</sup>13] Toru Tamaki, Junki Yoshimuta, Misato Kawakami, Bisser Raytchev, Kazufumi Kaneda, Shigeto Yoshida, Yoshito Takemura, Keiichi Onji, Rie Miyaki, and Shinji Tanaka. Computer-aided colorectal tumor classification in nbi endoscopy using local features. *Medical image analysis*, 17(1) :78 – 100, 2013.

- [VJ01] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages 1 – 9, 2001.
- [VJ04] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2) :137 – 154, 2004.
- [WBB<sup>+</sup>13] Amy Wang, Subhas Banerjee, Bradley A Barth, Yasser M Bhat, Shailendra Chauhan, Klaus T Gottlieb, Vani Konda, John T Maple, Faris Murad, Patrick R Pfau, et al. Wireless capsule endoscopy. *Gastrointestinal endoscopy*, 78(6) :805 – 815, 2013.
- [WBC04] Yu Wei, Xiong Bing, and Charayaphan Chareonsak. Fpga implementation of adaboost algorithm for detection of face biometrics. In *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, pages S1–6. IEEE, 2004.
- [WWL<sup>+</sup>12] Jie Wan, Ronggang Wang, Hao Lv, Lei Zhang, Wenmin Wang, Chenchen Gu, Quanzhan Zheng, and Wen Gao. Avs video decoding acceleration on arm cortex-a with neon. In *Signal Processing, Communication and Computing (ICSPCC), 2012 IEEE International Conference on*, pages 290–294. IEEE, 2012.
- [ZDA99] Gang Zhang, SG Demos, and RR Alfano. Far-red and nir spectral wing emission from tissues under 532 and 632 nm photo-excitation. *Lasers in the life sciences*, 9(1) :1–16, 1999.
- [ZDM<sup>+</sup>11] Qian Zhao, Themistocles Dassopoulos, Gerard E Mullin, MQ Meng, and Rajesh Kumar. A decision fusion strategy for polyp detection in capsule endoscopy. *Studies in health technology and informatics*, 173 :559 – 565, 2011.
- [ZFL10] Hongbin Zhu, Yi Fan, and Zhengrong Liang. Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps. *Beijing, China*, page 19, 2010.
- [ZIN14] Xin Zhou, Yasuaki Ito, and Koji Nakano. An efficient implementation of the one-dimensional hough transform algorithm for circle detection on the fpga. In *Computing and Networking (CANDAR), 2014 Second International Symposium on*, pages 447–452. IEEE, 2014.

# Annexe A

## Profils des vidéos et résultats des détections

Cette annexe montre les profils et les résultats des détections dans les 18 vidéos utilisées pour tester l'outil de détection des polypes.

Par profil de la vidéo, on sous entend si un polype est présent ou non dans l'image (c'est-à-dire la vérité terrain). Il s'agit de la courbe (a) sur toutes les figures (par exemple la figure A.1.a).

Les courbes (b), (c) et (d) représentent les résultats de la détection : on a 1 si le polype a été détecté (c'est-à-dire que le centre d'une des régions d'intérêt retourné pour l'outil tombe sur un polype) et 0 si le polype n'a pas été détecté. Les résultats sont obtenus en utilisant la méthode du centroïde.

La courbe (b) (par exemple la figure A.2.b) montre les résultats de la détection, sans aucun traitement dessus.

La courbe (c) (par exemple la figure A.3.c) montre les résultats de la détection après la fusion spatiale des différentes régions d'intérêt obtenues.

La courbe (d) (par exemple la figure A.4.d) montre les résultats de la détection après l'application de la cohérence temporelle.

Les caractéristiques d'apprentissage utilisées pour obtenir ces résultats sont les caractéristiques pseudo-Haar disposant d'une étape d'apprentissage actif.

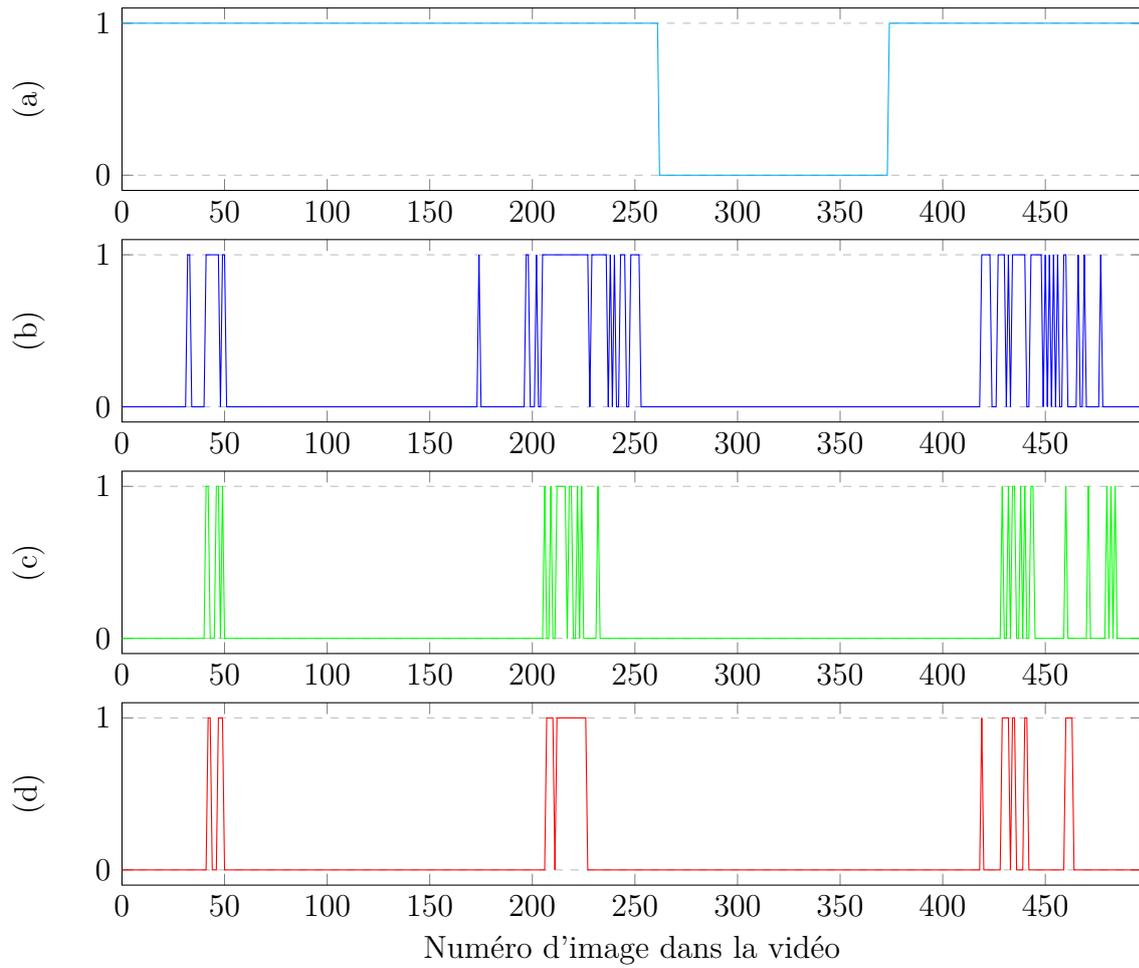


FIGURE A.1 – Résultats de la détection pour la vidéo 1 (498 images)

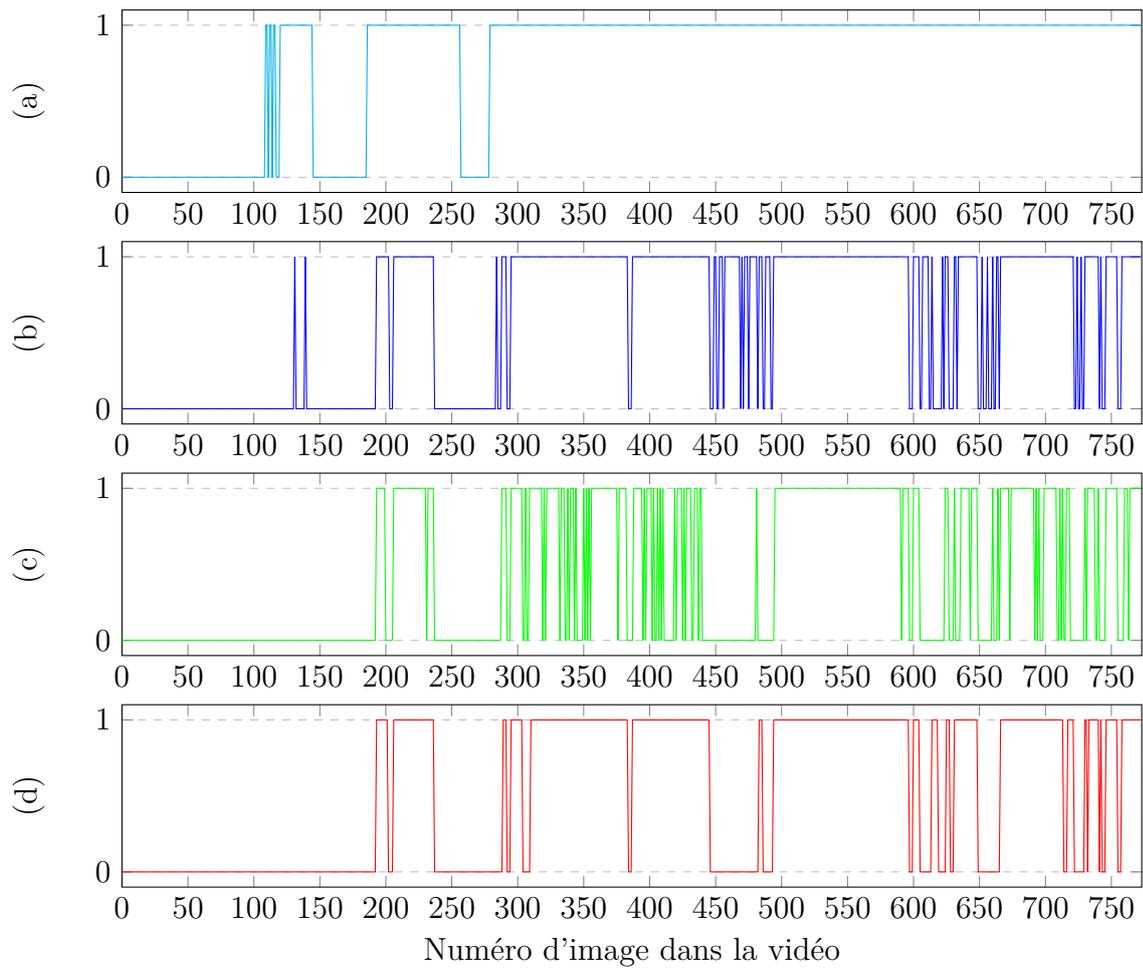


FIGURE A.2 – Résultats de la détection pour la vidéo 2 (773 images)

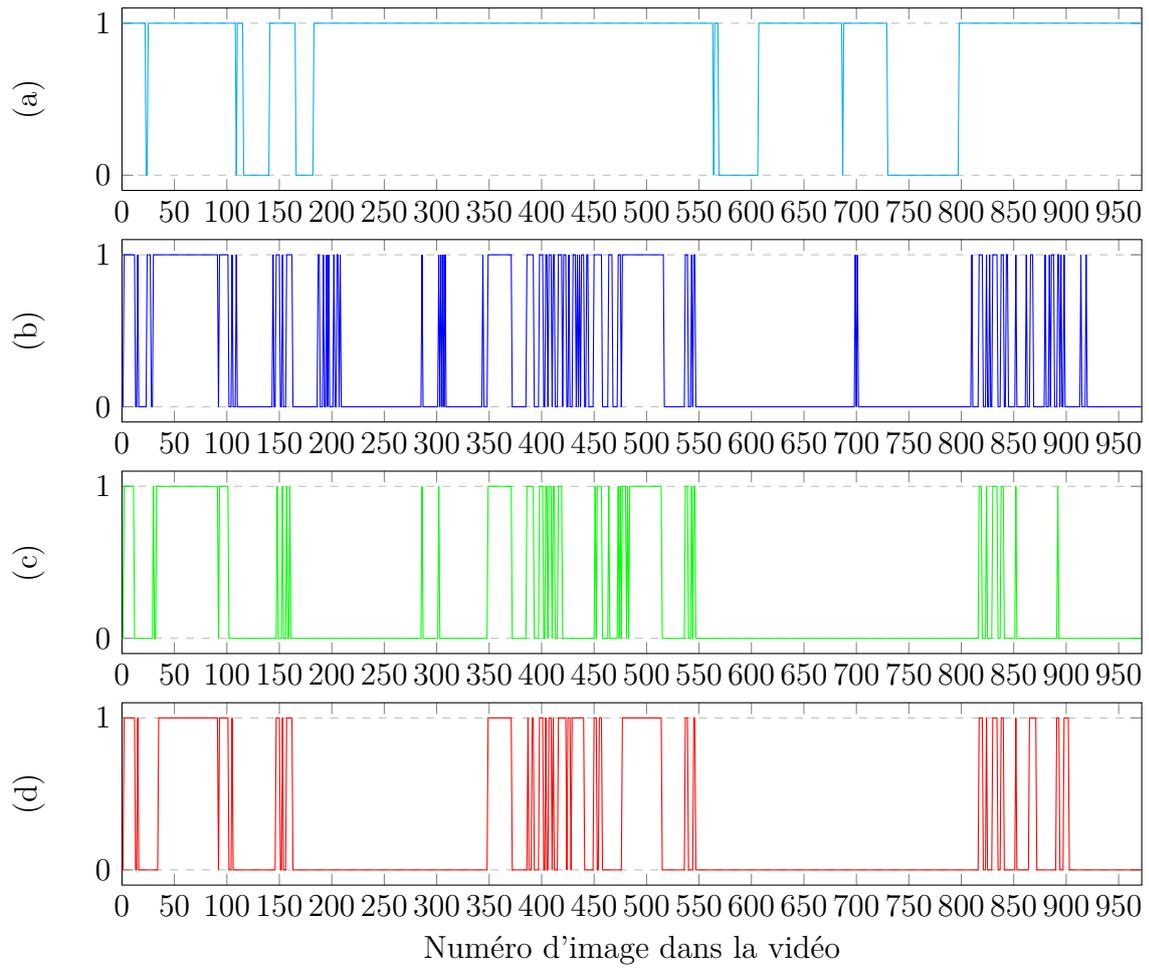


FIGURE A.3 – Résultats de la détection pour la vidéo 3 (972 images)

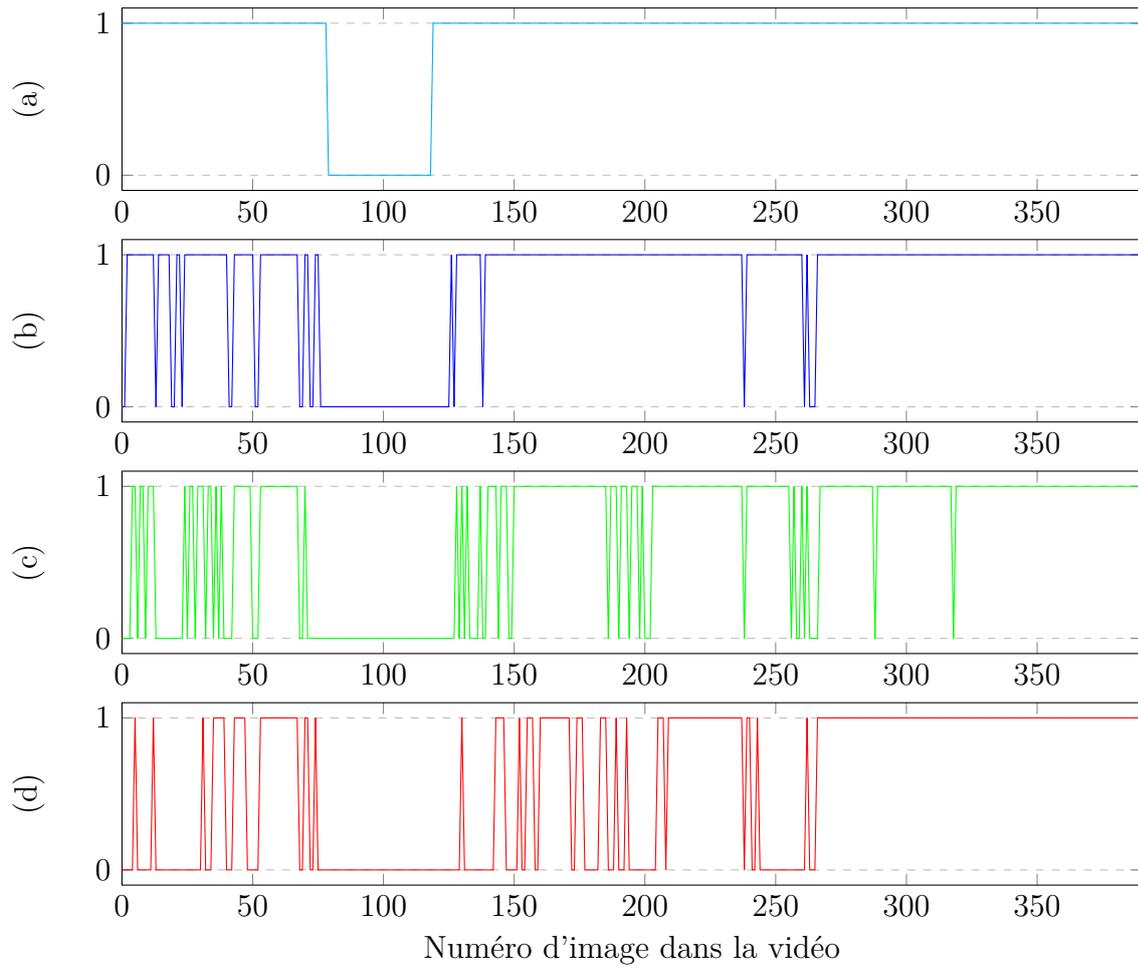


FIGURE A.4 – Résultats de la détection pour la vidéo 4 (390 images)

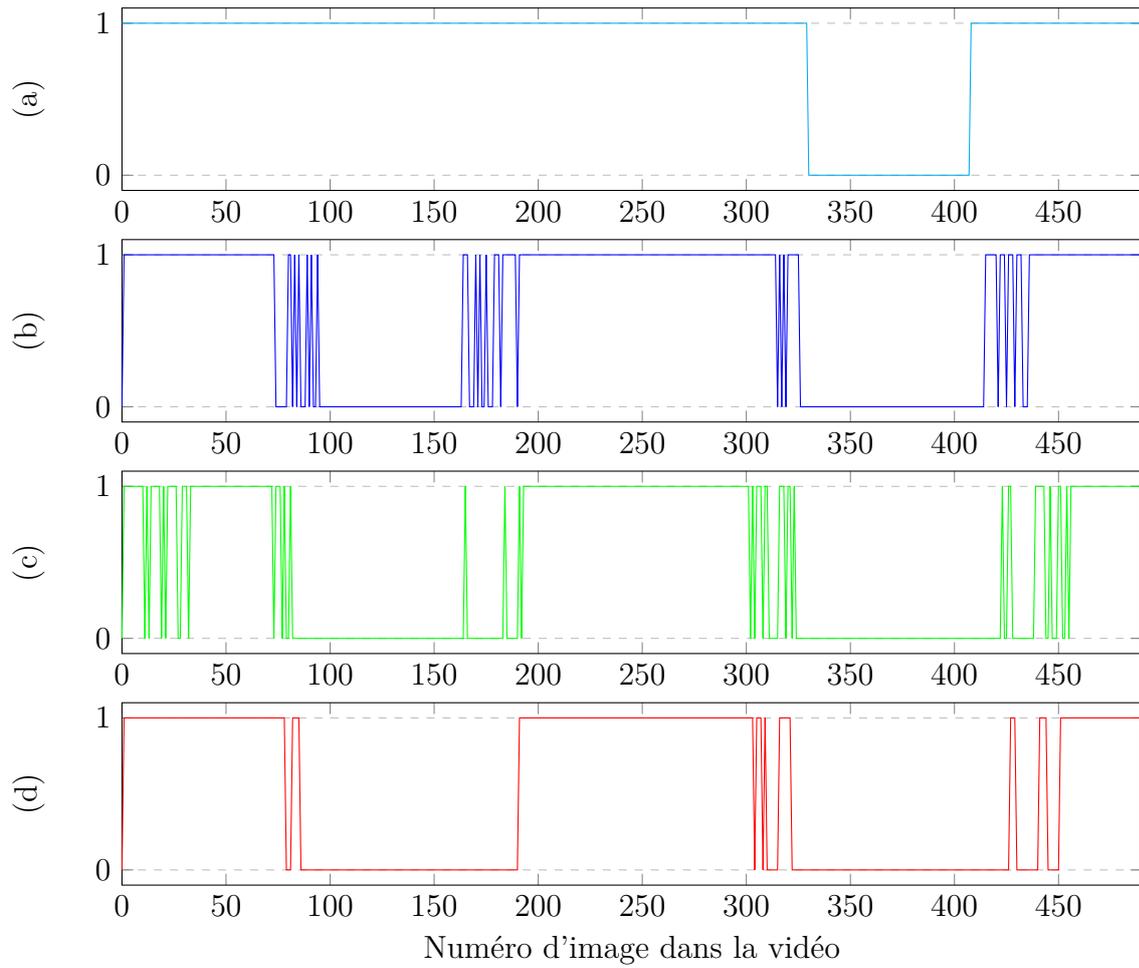


FIGURE A.5 – Résultats de la détection pour la vidéo 5 (490 images)

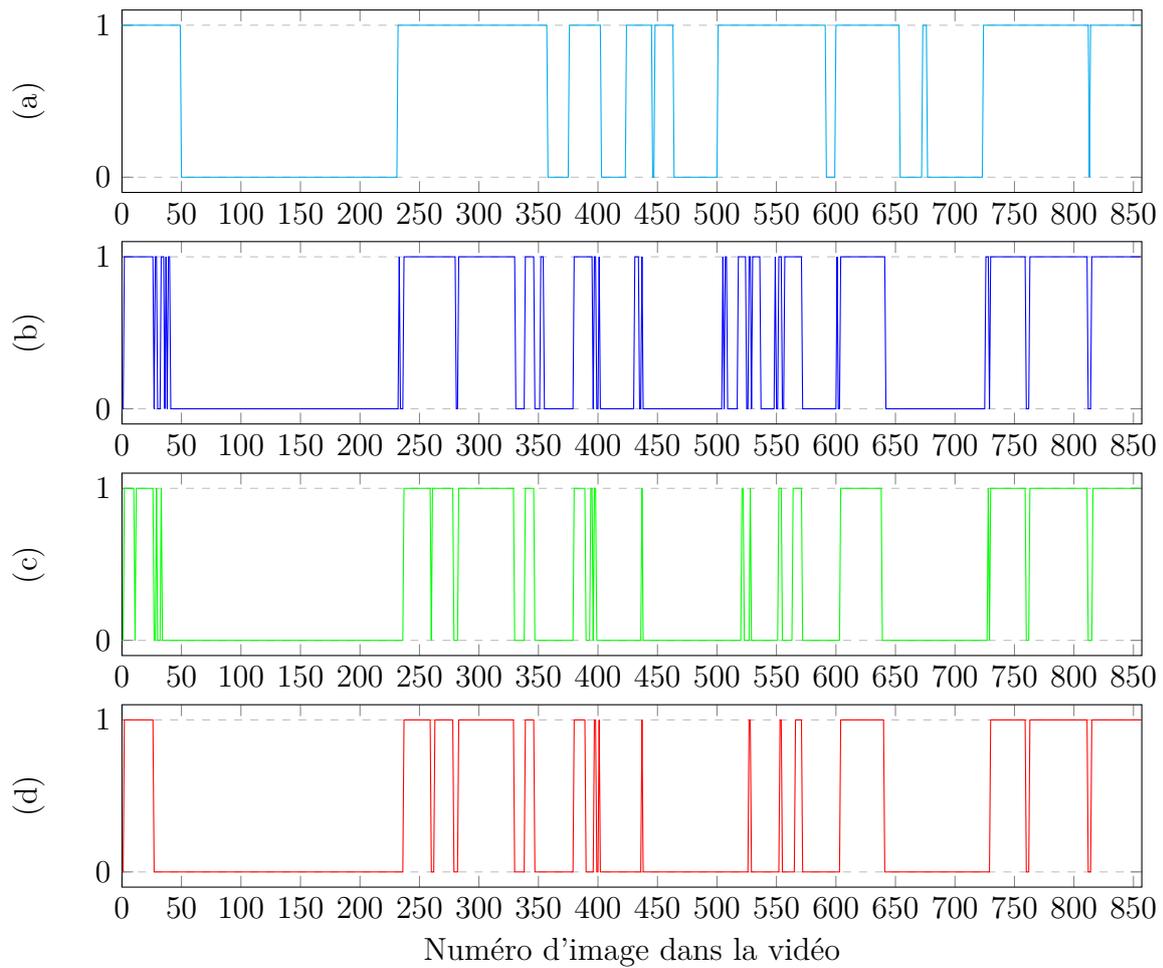


FIGURE A.6 – Résultats de la détection pour la vidéo 6 (857 images)

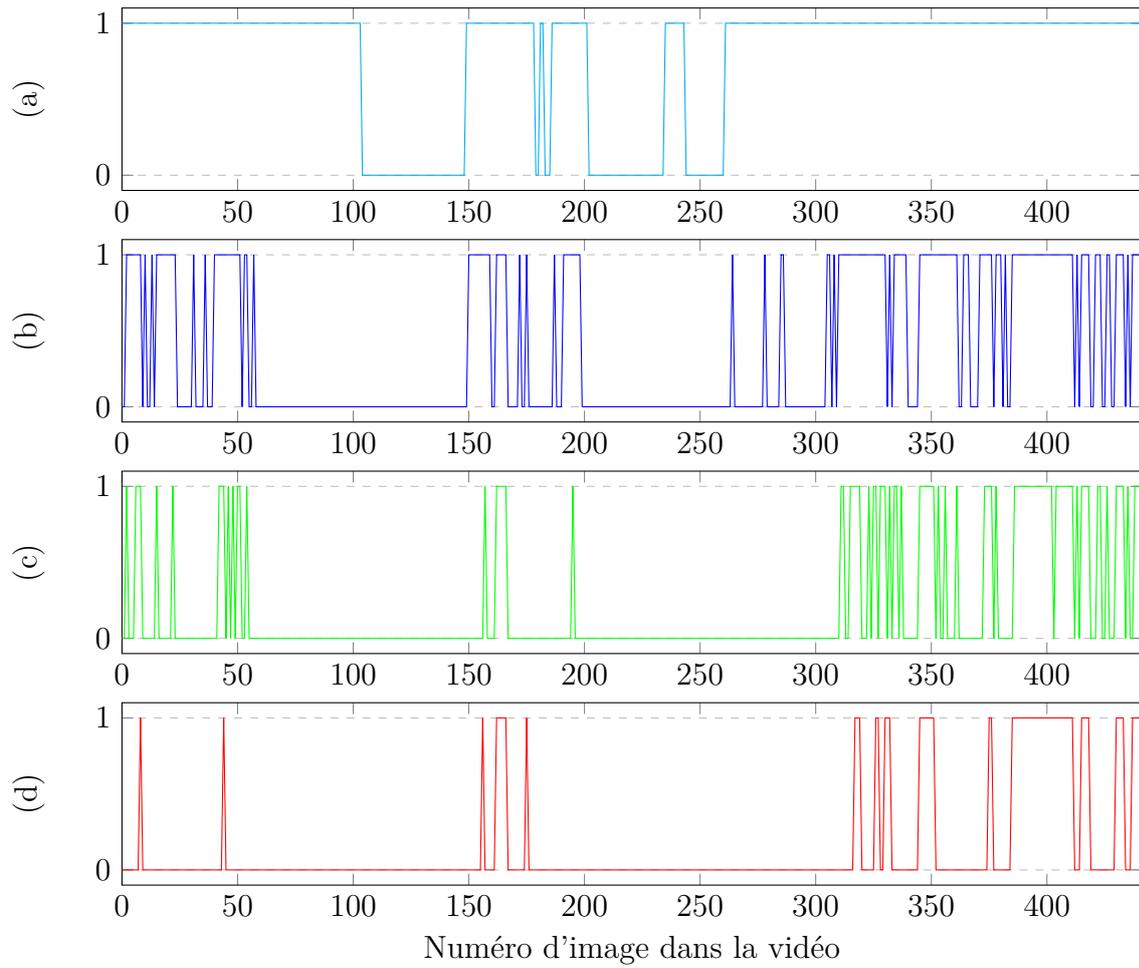


FIGURE A.7 – Résultats de la détection pour la vidéo 7 (441 images)

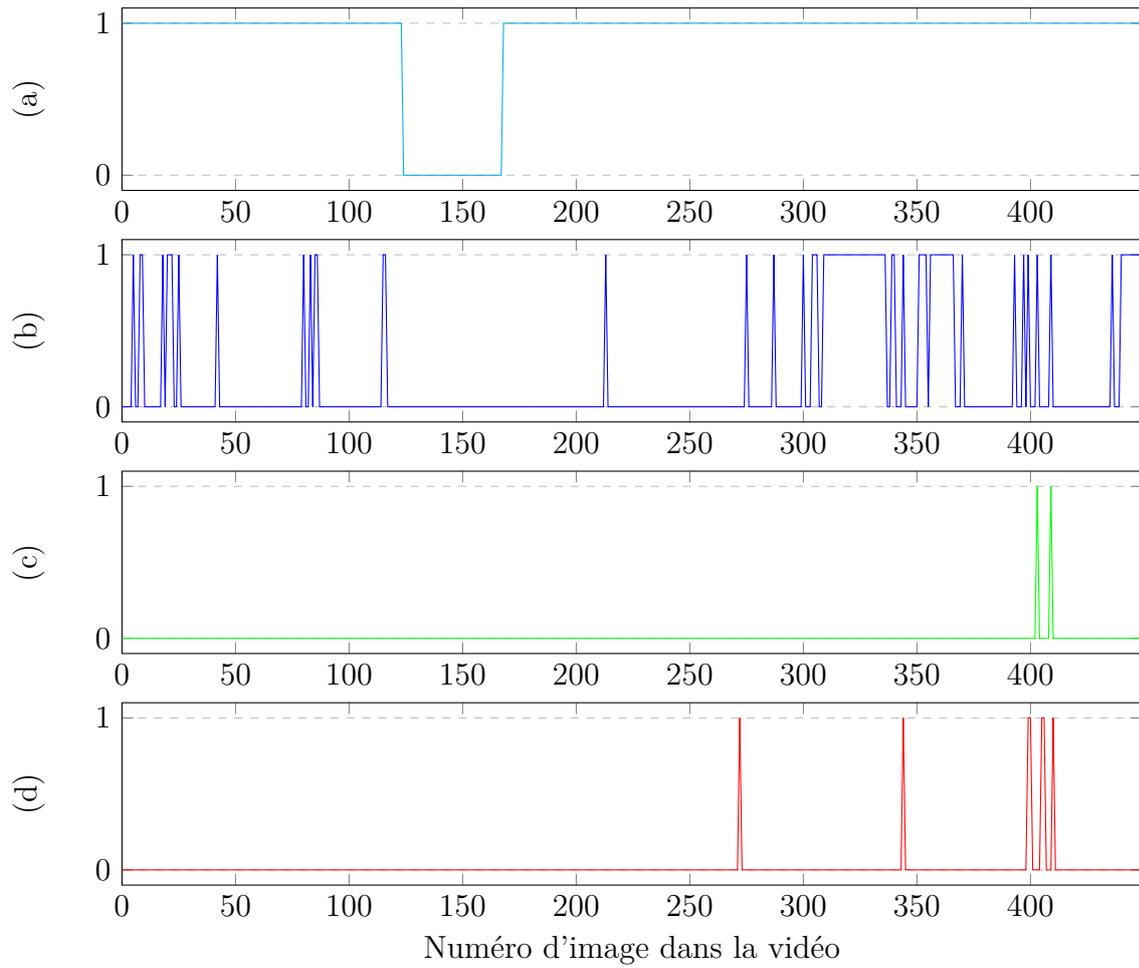


FIGURE A.8 – Résultats de la détection pour la vidéo 8 (449 images)

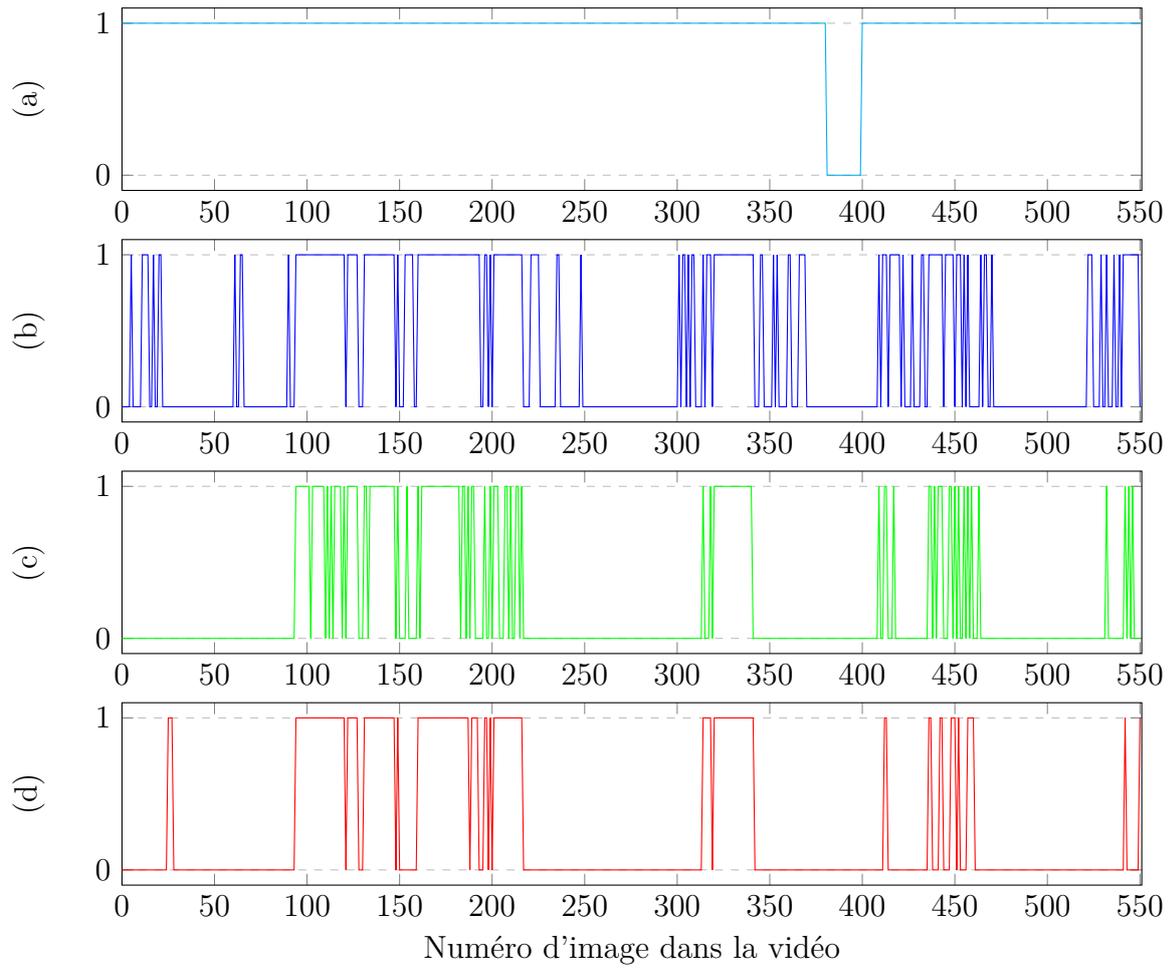


FIGURE A.9 – Résultats de la détection pour la vidéo 9 (551 images)

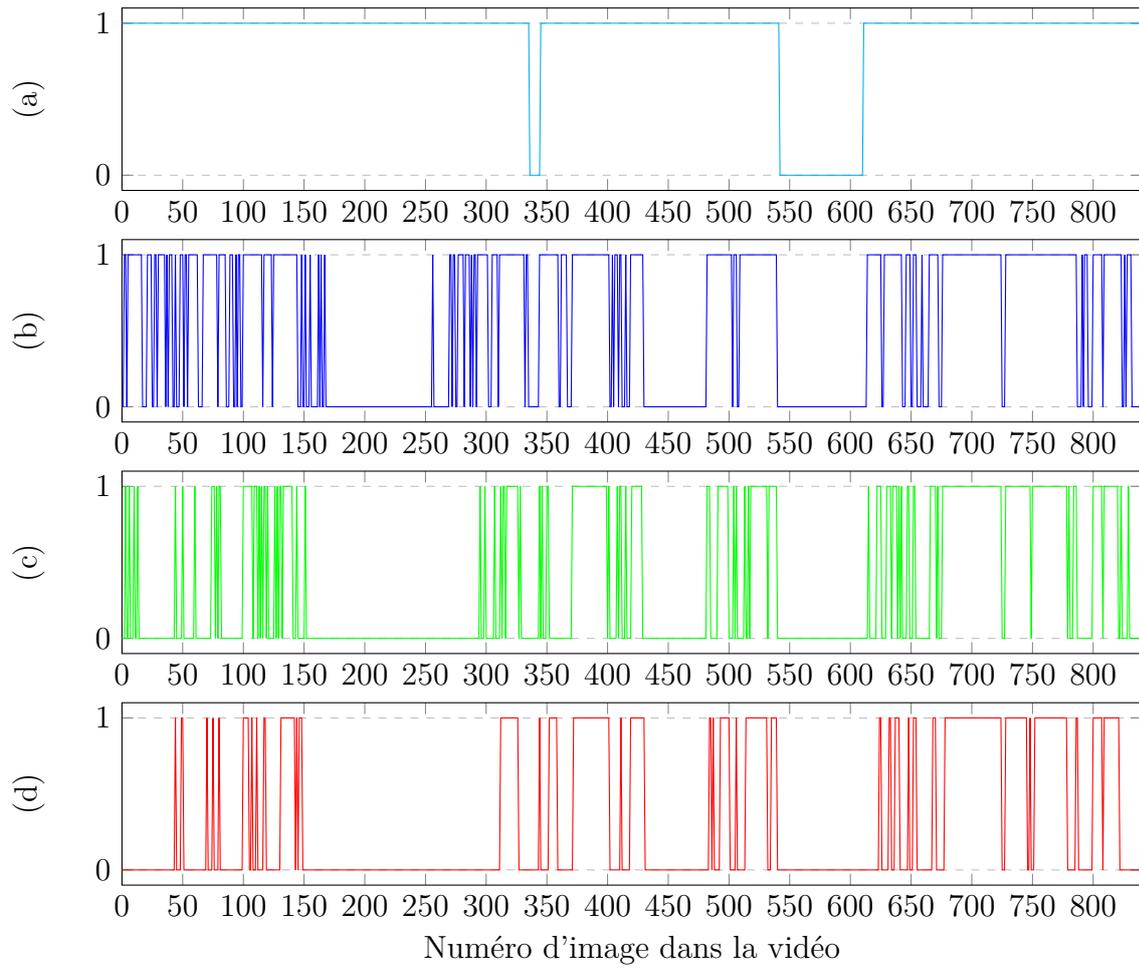


FIGURE A.10 – Résultats de la détection pour la vidéo 10 (840 images)

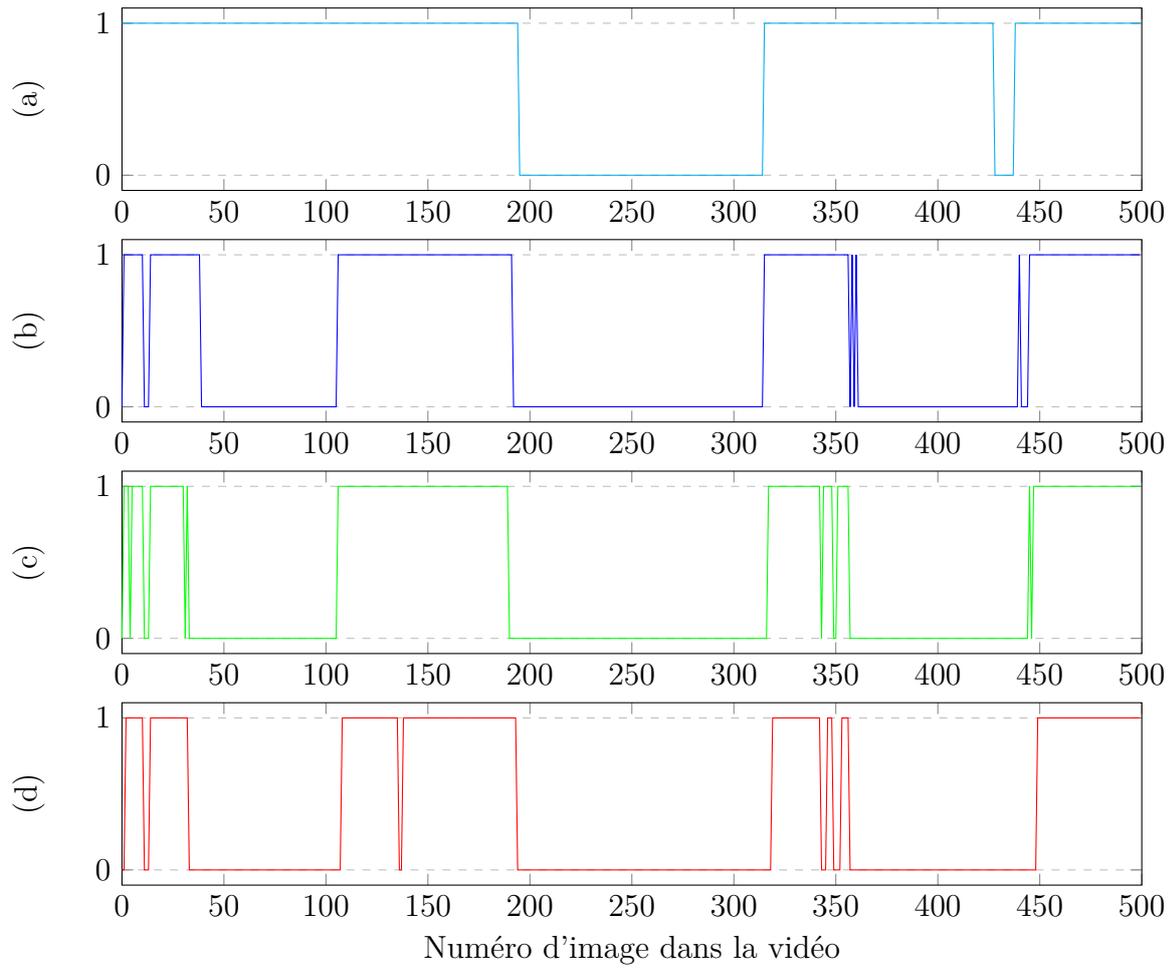


FIGURE A.11 – Résultats de la détection pour la vidéo 11 (500 images)

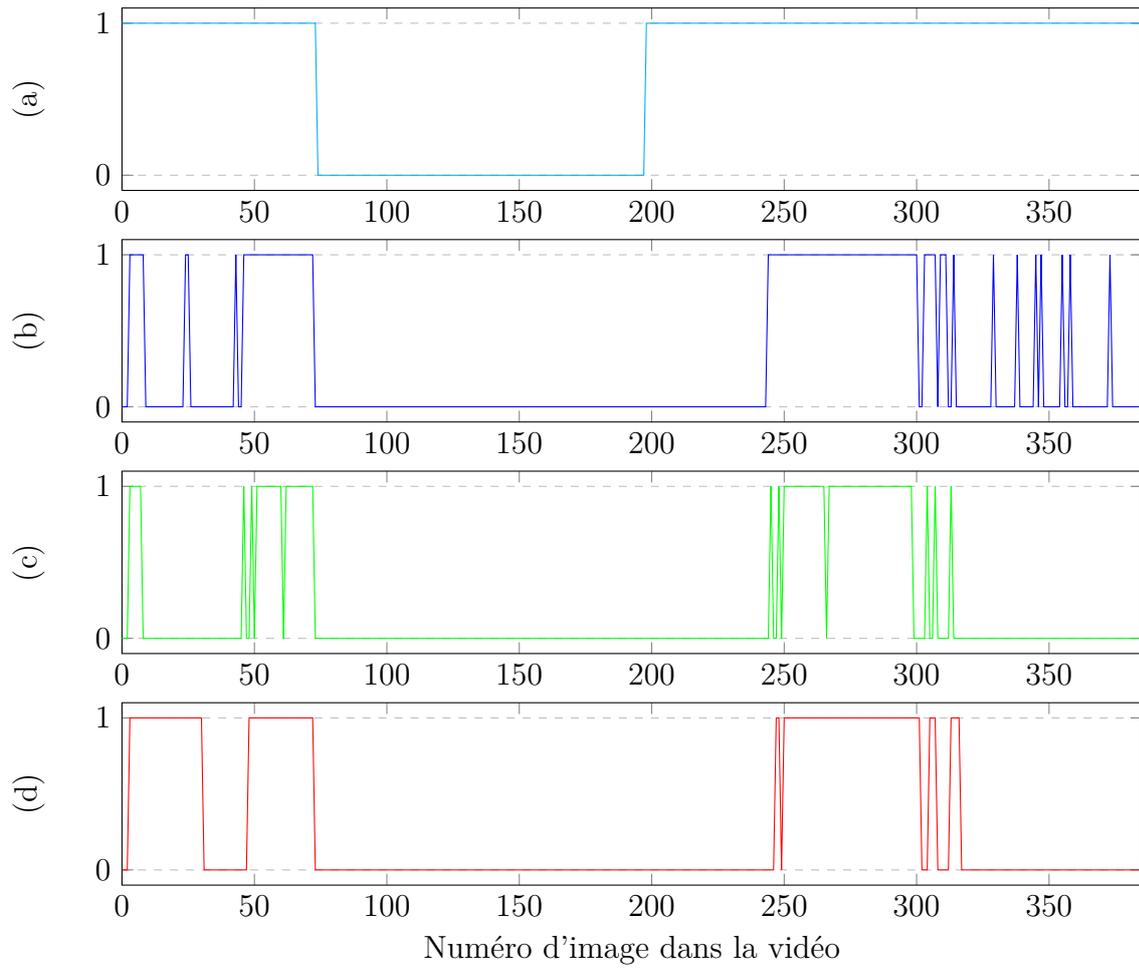


FIGURE A.12 – Résultats de la détection pour la vidéo 12 (385 images)

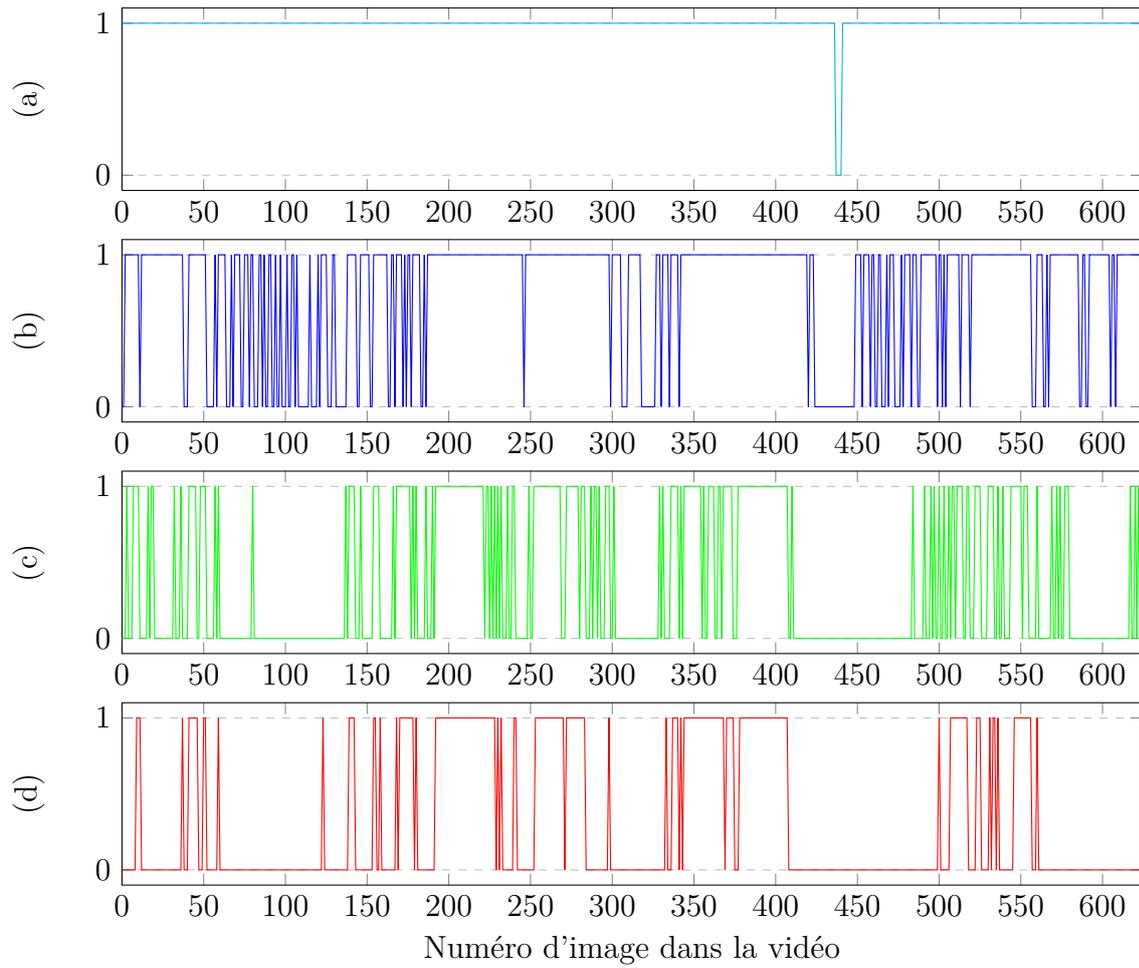


FIGURE A.13 – Résultats de la détection pour la vidéo 13 (624 images)

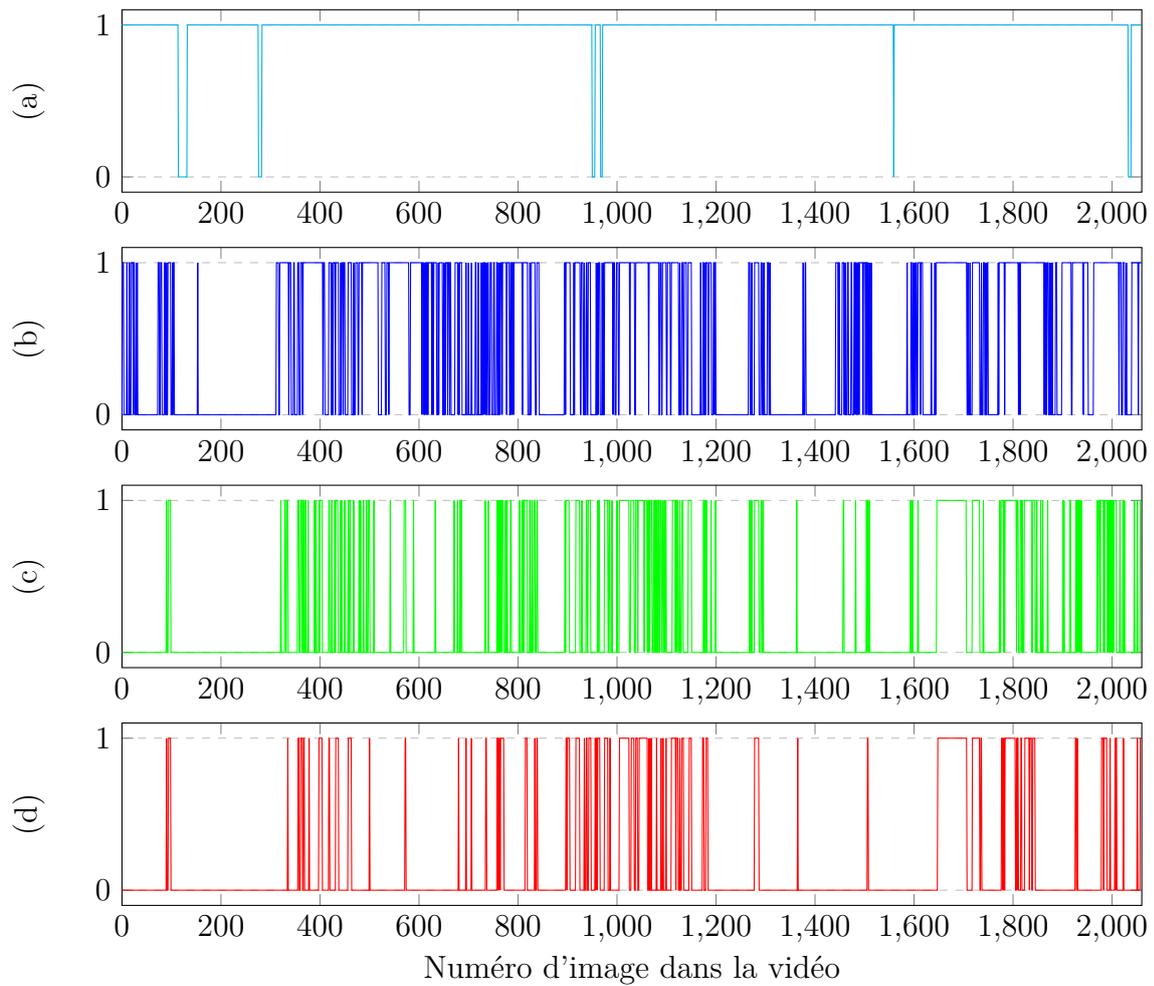


FIGURE A.14 – Résultats de la détection pour la vidéo 14 (2060 images)

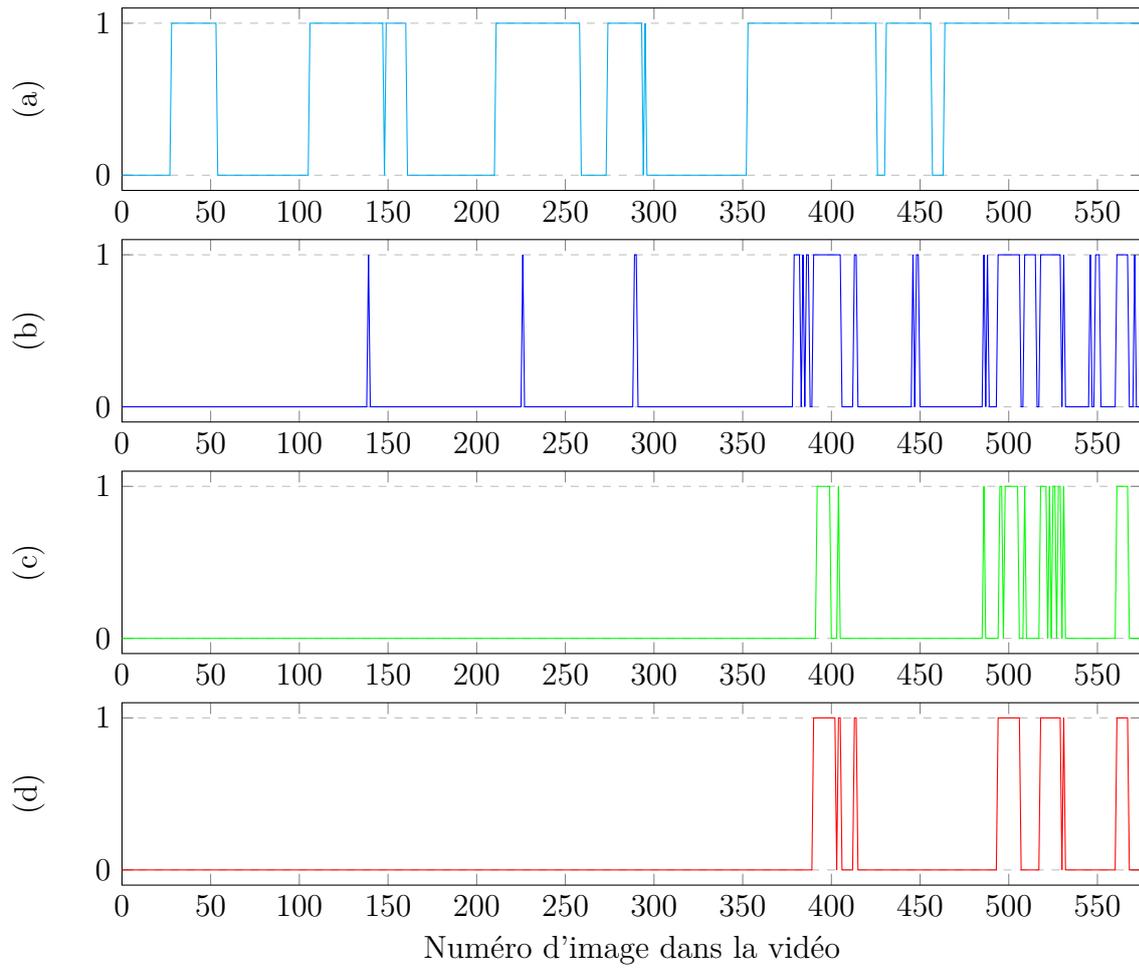


FIGURE A.15 – Résultats de la détection pour la vidéo 15 (575 images)

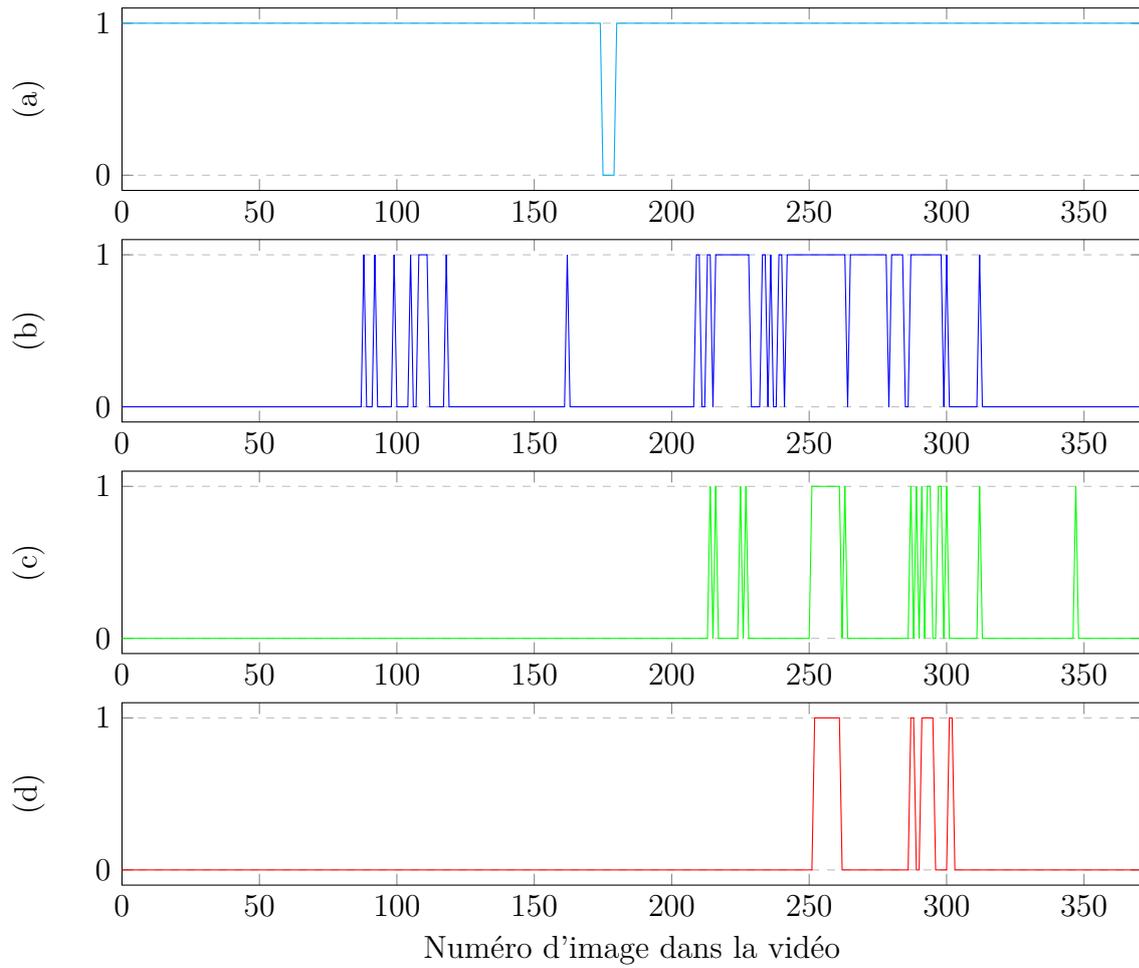


FIGURE A.16 – Résultats de la détection pour la vidéo 16 (371 images)

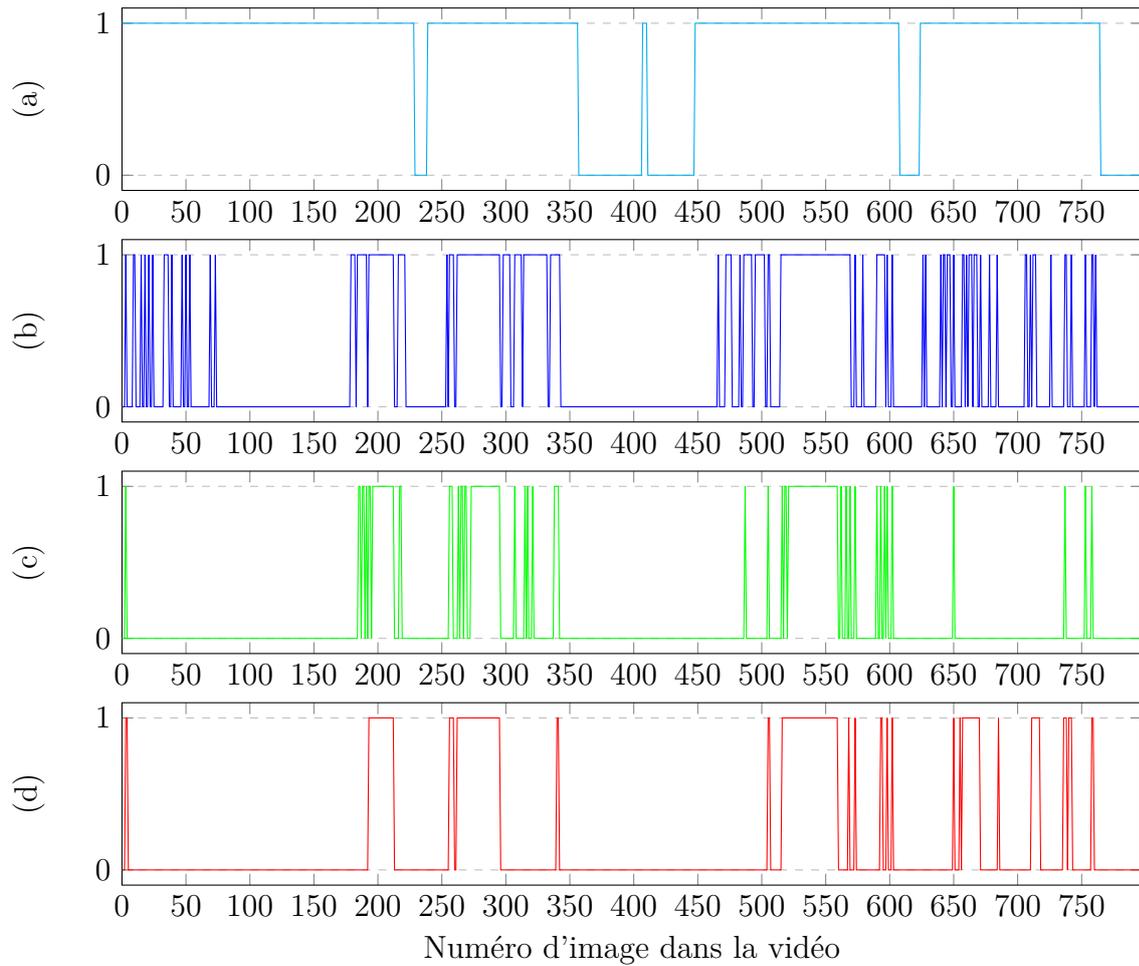


FIGURE A.17 – Résultats de la détection pour la vidéo 17 (797 images)

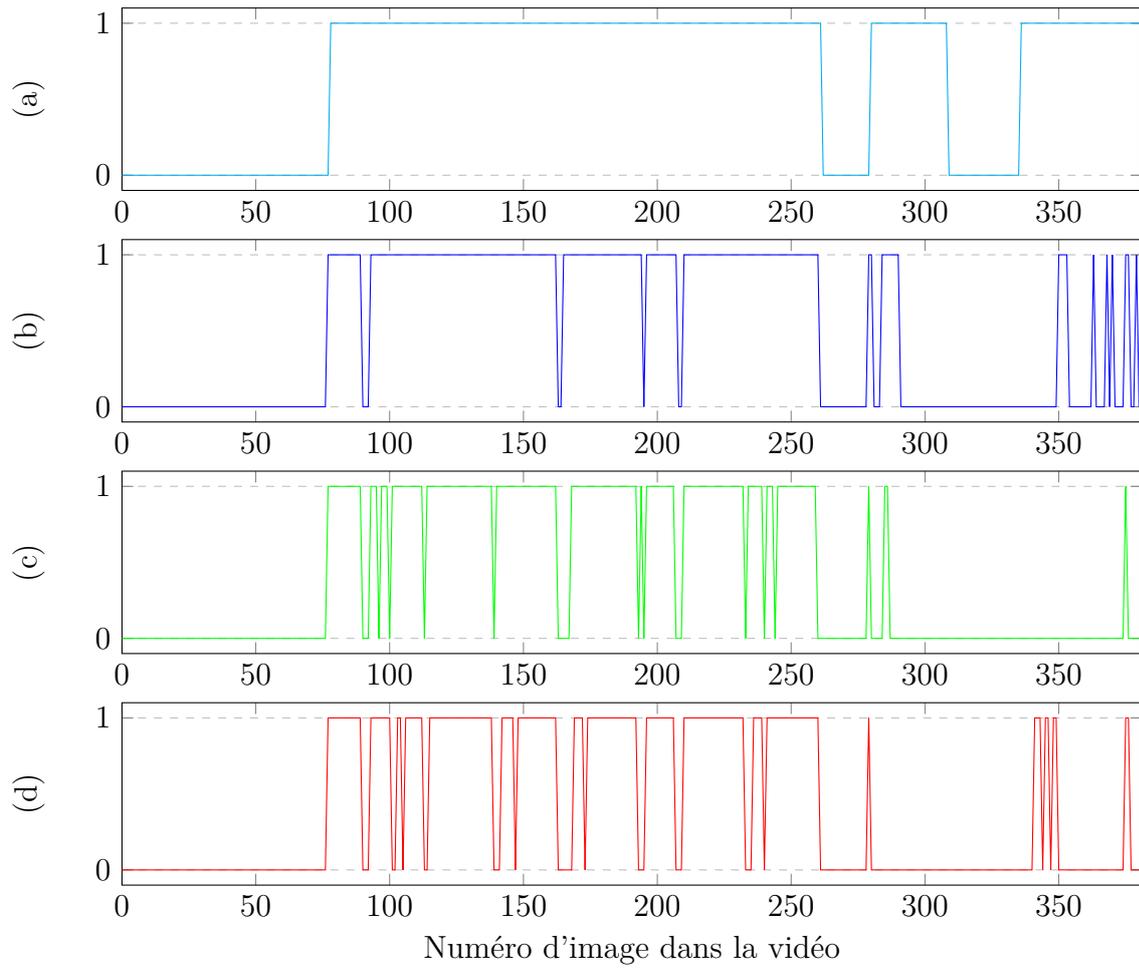


FIGURE A.18 – Résultats de la détection pour la vidéo 18 (381 images)



# Annexe B

## Implémentation matérielle, étude comparative

Cette annexe s'intéresse aux aspects d'implémentation matérielle de la méthode sur différentes plateformes afin d'en tester l'embarquabilité, leurs capacités à tenir le temps réel et leurs prix afin de proposer un système performant à prix abordable.

### B.1 De l'apprentissage vers l'implémentation matérielle

L'apprentissage pour la détection des polypes se passe à part. De ce fait, on n'a pas besoin de considérer son implémentation, la seule chose à savoir est que cette tâche est d'autant plus rapide que l'on dispose de puissance pour l'effectuer (dans notre cas, le temps d'apprentissage est monté jusqu'à 23 jours pour certains classificateurs). Juste pour information, cette tâche peut prendre entre quelques jours et quelques semaines. Une fois le classificateur créé, il suffit de le charger dans l'outil au début de l'exécution. Il est donc intéressant de mesurer le temps mis pour la détection sur différentes plateformes matérielles, qui possèdent des capacités plus ou moins accueillantes sur l'aspect embarquabilité, afin de les comparer et de déterminer quelle solution technique est la plus viable. On cherche donc ici à voir quel serait la plateforme qui possède le meilleur rapport performance - prix qui permettra de rendre intéressant notre outil pour les médecins.

### B.2 Embarquabilité

Dans ce chapitre, on s'intéresse à l'implémentation matérielle de la méthode sur différentes plateformes embarquables afin de trouver celle qui est la plus adaptée à notre problématique. Le but ici est de montrer quelles sont les solutions technologiques qui s'offrent à nous et lesquelles sont les plus intéressantes pour accueillir notre outil de détection des polypes.

Dans la littérature, en faisant un tour rapide des méthodes existantes, à notre connaissance (tableau B.1), aucune de ces méthodes n'est capable de traiter les images de colo-

| Auteurs                             | Compatible temps réel   |
|-------------------------------------|---|
| Li et al. [LM12]                    | Non défini  |
| Tamaki et al. [TYK <sup>+</sup> 13] | Non (0.60 secondes par image, avec une résolution entre 100x300 et 900x800) |
| Mamonov et al. [MFFT14]             | Non   |
| Zhao et al. [ZDM <sup>+</sup> 11]   | Non défini  |
| Bernal et al. [BSV12]               | Non (19 secondes par image, avec une résolution de 574x500)                 |
| Angermann et al. [AHR16]            | Oui (39 millisecondes par image, avec une résolution de 574x500)            |

TABLE B.1 – Compatibilité avec le temps réel des systèmes dans la littérature

scopie en temps réel afin de retourner la localisation du polype dans l’image au médecin. Si on veut assurer une utilisation de notre méthode dans la routine clinique quotidienne, il faut donc s’intéresser à cette contrainte de temps réel, qui va permettre d’assister le médecin dans son examen sans le perturber avec un certain délai entre ce que voit le médecin à l’endoscope et le résultat de la détection.

C’est pour cette principale raison que le développement d’un nouveau système capable de détecter automatiquement les polypes sans altérer le processus de visualisation du médecin lors de l’examen est un besoin indéniable. Ce système doit donc pouvoir aider le médecin en fournissant en temps réel la détection des lésions avec deux objectifs principaux : le premier, réduire le taux de polypes manqués et le second, intégrer tous les traitements requis sur une plate-forme matérielle compatible avec une utilisation facile dans les salles d’examen.

Pour la comparaison des différentes plateformes, on se propose d’utiliser la méthode développée permettant de détecter les polypes dans des images fixes issues de vidéocoloscopie. Dès le début du développement de notre méthode, on a porté une attention particulière à la compatibilité de notre méthode avec des systèmes embarqués, et en utilisant des fonctions compatibles avec ce but.

Notre outil, une fois porté sur plateforme, se doit d’être efficient en terme de capacité de détection, fonctionnant en temps réel, mais aussi être facile d’accès pour les médecins et facilement utilisable en salle d’examen. Selon [MCE<sup>+</sup>16], dans les salles d’examens, le matériel disponible est basé sur les processeurs suivant : Intel Core2Duo, Intel i5 ou Intel Pentium Dual Core qui peuvent fonctionner à différentes fréquences. C’est pour cela qu’on a décidé, pour tester notre outil, de prendre des plateformes basées sur le même type de processeur, mais aussi d’autres types. On a donc décidé de comparer plusieurs plateformes, plus ou moins embarquables et de différents types. Les plateformes testées sont les suivantes (le choix à été restreint par leurs disponibilités) :

- Lenovo Yoga 2 Pro : il s’agit d’un ordinateur portable grand public avec un processeur Intel Core i5 (quad-core @ 2.3 GHz) et disposant de 4Go de mémoire RAM

tournant sous Windows 10.

- Dell Precision : il s'agit d'une station de calcul avec un processeur Intel Xeon E5 (hexa-core @ 2.8 GHz) et disposant de 32Go de mémoire RAM tournant sous Windows 8.1.
- Raspberry Pi2 : il s'agit d'un petit ordinateur avec un processeur ARM Cortex-A7 (dual-core @ 900 MHz) et disposant de 1Go de mémoire RAM tournant sous Raspbian Jessie 4.4.
- Raspberry Pi3 : il s'agit d'un petit ordinateur avec un processeur ARM Cortex-A53 (quad-core @ 1.2 GHz) et disposant de 1Go de mémoire RAM tournant sous Raspbian Jessie 4.4.
- nVidia Jetson TK1 : il s'agit d'un kit de développement de nVidia contenant un CPU (ARM Cortex-A15, quad-core @ 2.3 GHz), un GPU (architecture nVidia Kepler avec 192 coeurs CUDA) et disposant de 2Go de mémoire RAM tournant sous "Linux For Tegra" (Ubuntu 16.04). (On utilise sur cette plateforme à la fois le CPU et le GPU)
- Altera DE1 SoC Board : il s'agit d'un kit de développement d'Altera based sur un FPGA "System-On-Chip" avec un processeur ARM Cortex-A9 (dual-core @ 800 MHz) et disposant de 1Go de mémoire RAM et tournant sous Ubuntu 16.0. (On utilise sur cette plateforme uniquement le processeur)

Sur chacune des plateformes est installée une des dernières versions d'OpenCV (qui est la 3.1.0). Pour rappel, OpenCV est une librairie libre de droits disposant de plusieurs fonctions dédiées au traitement d'images, au traitement de vidéos, à l'apprentissage machine et bien d'autres. De plus, OpenCV peut être utilisé pour réaliser des interfaces graphiques ou pour faire de la réalité augmentée. Notre utilisation d'OpenCV s'arrête ici à utiliser uniquement les fonctions de traitement d'images et d'apprentissage machine.

Pour comparer les performances des différentes plateformes, on se propose d'utiliser une parties des métriques déjà présentées :

- Les vrais positifs (True Positive, TP). Une détection "vraie positive" est le résultat lorsqu'un polype est détecté correctement.
- Les faux positifs (False Positive, FP). Une détection "fausse positive" est le résultat lorsqu'un polype est signalé négatif alors qu'il devrait être positif.
- Les faux négatifs (False Negative, FN). Une détection "fausse négative" est le résultat lorsqu'un polype est signalé positif alors qu'il n'y a rien.
- Le rappel ( $= \frac{TP}{TP+FN}$ ). Le rappel permet de définir le résultat de bonnes détections par rapport au nombre total de lésions à détecter.
- La précision ( $= \frac{TP}{TP+FP}$ ). La précision permet de définir le résultats de bonnes détections par rapport au nombre total de lésions détectées.
- Le F<sub>2</sub> Score ( $= \frac{5 \times \text{Précision} \times \text{Rappel}}{4 \times \text{Précision} + \text{Rappel}}$ ). Le F<sub>2</sub> Score, qui combine Précision et Rappel, est utilisé pour faire la balance entre le nombre de fausses alarmes et le nombre de lésions ratées.
- Le temps de détection moyen pour une image. Il s'agit du temps moyen (moyenne prise sur l'ensemble des images testées) mis par la méthode pour effectuer la détection des polypes dans une image.

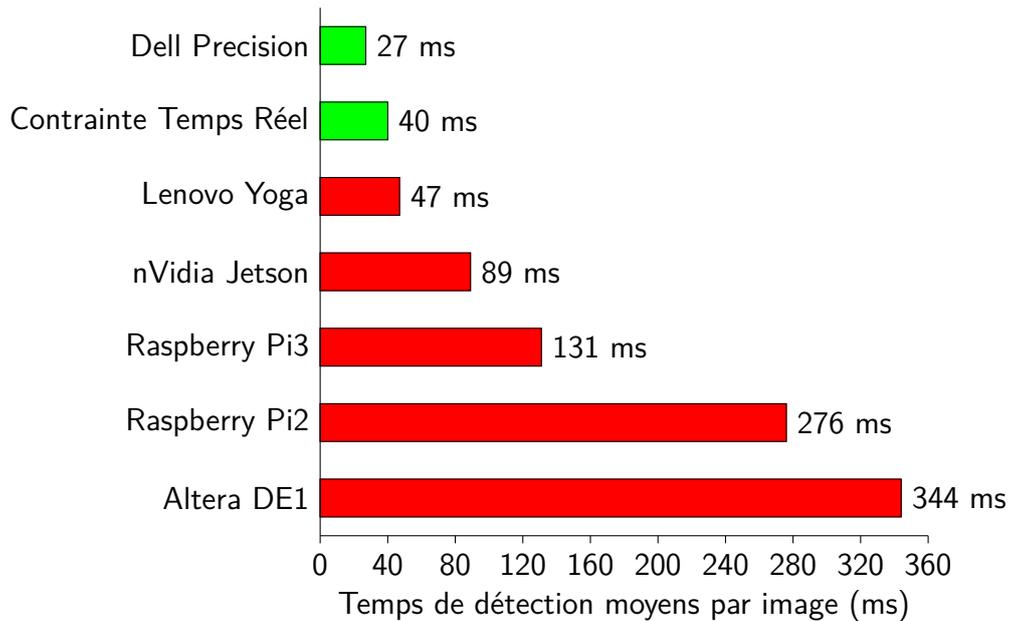


FIGURE B.1 – Temps de détection moyens par image pour chacune des plateformes

Les tests sont menés sur le set d'images CVC-ColonDB [BSV12] qui contient seulement 273 images exploitables (sur les 300 présentes). Ces images possèdent une résolution de 574x500 pixels.

### B.2.1 Temps de détection moyen sur différentes plateformes

Le premier test a consisté à comparer les temps de détection moyens par image sur chacune des plateformes, et en particulier de voir si ces plateformes sont capables de traiter les images en temps réel. Par traiter les images en temps réel, on sous-entend que le temps de calcul doit être inférieur ou égal à 40 millisecondes afin de ne pas perturber le médecin dans son examen et ne pas ajouter de retard entre ce que voit le médecin et le résultat de notre détection. Les résultats sont disponibles sur la figure B.1.

Quand on regarde les résultats de la détection (figure B.1), on s'aperçoit que la station de calcul est la seule plateforme capable actuellement de faire la détection en temps réel. Mais on voit aussi qu'un ordinateur portable grand public est proche aussi du temps réel, et si on prend en compte l'évolution des performances des matériels informatiques et l'âge de l'ordinateur, on peut sans aucun doute dire qu'un ordinateur grand public actuel sera capable d'effectuer la détection en temps réel. Concernant les autres plateformes, on est clairement loin des performances de détection en temps réel mais il existe des optimisations pour palier à ce problème.

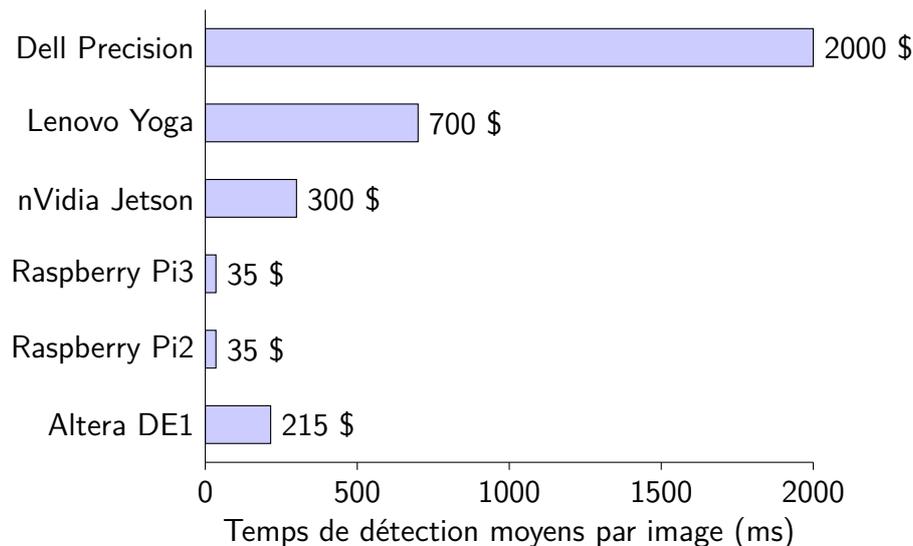


FIGURE B.2 – Rapport performances - prix des différentes plateformes

### B.2.2 Coût des plateformes

Afin de proposer une plateforme à prix abordable, on s'intéresse aussi aux prix des différentes plateformes testées afin de déterminer laquelle possède le meilleur ratio prix - capacité temps réel. Les prix d'achat des différentes plateformes sont disponibles dans la figure B.2.

Grâce à la figure B.2, on peut s'apercevoir que c'est la Raspberry Pi3 qui tire son épingle du jeu. En effet, avec un prix de 35 \$ et un temps de calcul de 131 millisecondes pour traiter une image, c'est elle qui semble être la meilleure pour effectuer la détection embarquée des polypes dans les salles d'examen. C'est donc la Raspberry Pi3 que l'on choisira comme candidate.

### B.2.3 Cas des processeurs ARM : accélération

Concernant le cas particulier des processeurs ARM (comme la Raspberry Pi3), le fabricant propose un jeu d'instructions SIMD (Single instruction multiple data, ou instruction unique, données multiples en français) qui permet de paralléliser le traitement afin d'optimiser et d'améliorer les temps de calcul pour le traitement multimédia sur leurs processeurs. Cette fonctionnalité s'appelle NEON. Cela va donc nous permettre d'améliorer notre temps de calcul sur processeur ARM. Dans la littérature, certaines publications [MJR<sup>+</sup>13] [WWL<sup>+</sup>12] ont montré une amélioration notable du temps de calcul pour le traitement des données multimédia. Pour notre application, on a décidé de comparer les performances sur un processeur ARM dont on disposait. Pour des raisons pratiques, cette comparaison s'est faite sur la carte de développement Altera DE1 SoC. Le test a été effectué en utilisant la même version d'OpenCV, une avec l'option NEON activée, et l'autre sans. Comme précédemment, on a utilisé la même base de données d'images qui était la CVC-ColonDB [BSV12]. Les résultats des temps de calculs sont disponibles sur la figure

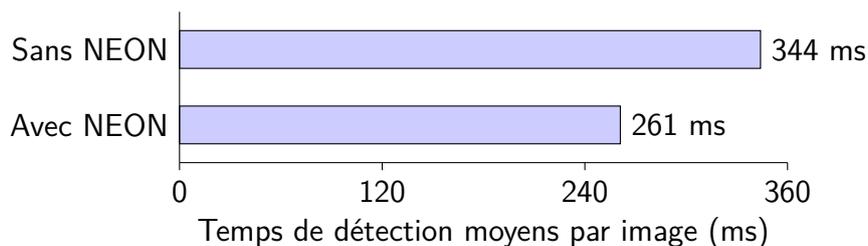


FIGURE B.3 – Temps de détection avec et sans NEON sur Altera DE1 SoC

### B.3.

Comme on peut le constater sur la figure B.3, le temps de détection sur la carte Altera passe de 344 millisecondes à 261 millisecondes, soit une réduction d'environ 24% du temps alloué à la détection des polypes dans les images de coloscopie. Le résultat constaté est aussi en accord avec certains résultats de la littérature qui annonce des facteurs d'accélération compris entre 1,05 et 13,88 [MJR<sup>+</sup>13] [CP11]. Cela permet donc d'envisager de porter sur plateforme ARM notre outil de détection dont les performances vont se rapprocher de plus en plus vers le temps réel grâce aux optimisations disponibles.

## B.3 Conclusion

On a donc pu voir ici différentes plateformes envisagées pour accueillir notre méthode de détection des polypes. Parmi toutes les méthodes proposées, on a mesuré les temps de calculs moyens pour traiter une image. On a donc pu voir sans surprise que les stations de calculs s'en sortaient mieux que les autres. Mais toutes les plateformes testées ne sont pas forcément inutiles. Comme on souhaite proposer un outil abordable, on a aussi comparé les coûts de ces différentes plateformes. On a ainsi pu voir qu'au final, au regard du coût et des performances, que la Raspberry Pi3 était la plateforme la plus apte à recevoir notre méthode. De plus, il existe des optimisations pour traiter les médias sur les processeurs de type ARM, qui permettent d'avoir un gain de 24% en temps de calculs. Ce qui permet de confirmer le choix de la Raspberry Pi3 en tant que plateforme adéquate pour notre outil. Cependant, il faut trouver de nouvelles optimisations afin de réduire le temps de calculs.

# Annexe C

## Détection des naevus

Dans cette annexe, on montre l’adaptabilité de notre méthode à un autre problème d’ordre médical : la détection et la caractérisation automatique des naevus.

### C.1 Introduction

Le cancer de la peau est un des cancers les plus diagnostiqué dans le monde [FSE<sup>+</sup>15]. Ce cancer se base sur deux types de lésions : les carcinomes et les mélanomes. Les deux sont reliés à la dégénération d’un naevus. Les carcinomes peuvent causer des cancer locaux mais provoquent rarement des métastases. Ce sont les mélanomes qui sont responsables principalement de ce cancer et des décès associés. Le moyen le plus efficace de lutter contre ce cancer est de pratiquer, tout comme pour le cancer colorectal, des examens de dépistages précoces et de retirer les lésions suspectes. Pour déterminer si un naevus est dangereux, le médecin applique la méthode ABCDE [MAAJ13] en regardant simplement le naevus :

- A pour l’Asymétrie du naevus
- B pour l’irrégularité des Bordures du naevus
- C pour l’homogénéité de la Couleur du naevus
- D pour le Diamètre du naevus
- E pour l’Evolution du naevus dans le temps

L’intérêt de ce problème (la détection et la caractérisation des naevus) est de savoir si la méthode de détection des polypes est adaptable à un autre problème dont le but est totalement différent mais la méthode pour y arriver est la même. C’est pour cela qu’on s’est intéressé à ce problème. Pour y arriver, on a développé la chaine de traitement présentée en figure C.1 [AHR17]. Cette chaine de traitement se décompose en 4 étapes distinctes (tableau C.1)

- 
1. Une image contenant un (ou plusieurs) naevus est présentée en entrée
  2. S'en suit alors une étape de détection (basée sur la même méthode que pour les polypes) et retourne des régions d'intérêt contenant un possible naevus
  3. Puis chaque région d'intérêt est caractérisée (c'est-à-dire savoir si le naevus détecté est bénin ou malin) en appliquant les critères ABCD
  4. Et le résultats est retourné à l'utilisateur sous forme d'un indice de confiance entre 0% et 100% (Le minimum signifiant un naevus malin, et le maximum un naevus bénin)
- 

TABLE C.1 – Méthode de détection des naevus

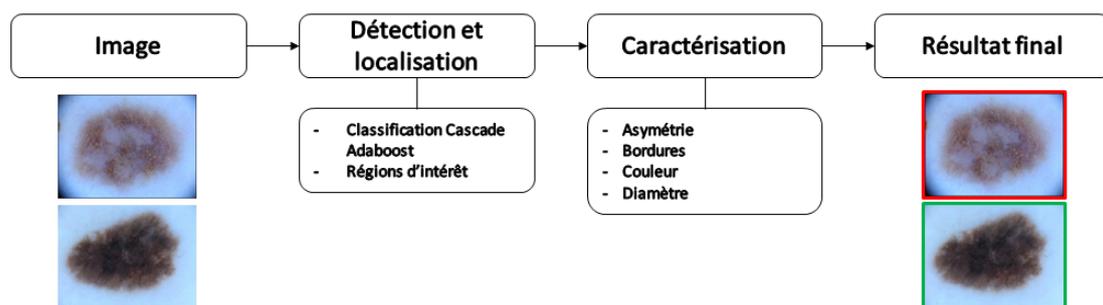


FIGURE C.1 – Chaîne de traitement des naevus

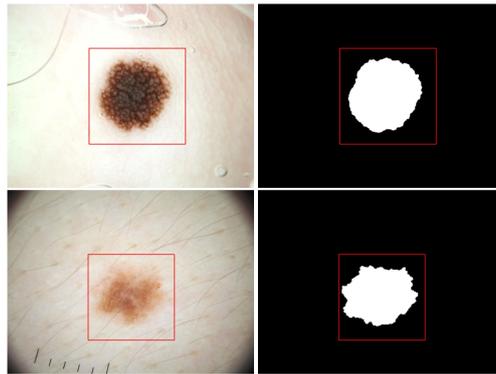


FIGURE C.2 – Exemple de détections et comparaison à la vérité terrain

## C.2 Détection et localisation des naevus

Afin de détecter les naevus, on utilise donc l'apprentissage machine et en particulier la méthode de Cascade Adaboost [VJ01] combinée à notre méthode de détection des polypes [ABSM<sup>+</sup>17b]. Il est aussi tout à fait possible d'utiliser l'apprentissage actif afin d'améliorer la classification quand la base de données est pauvre en images.

Concernant l'apprentissage, il est effectué sur la base de données mise à disposition dans le challenge ISBI [GCC<sup>+</sup>16]. Sur la totalité des images disponibles (900 images), on en a conservé 600 pour apprendre et 300 pour effectuer des tests.

Grâce à cette technique, on est alors capable de renvoyer des régions d'intérêt contenant des possibles naevus afin de les caractériser. Quelques exemples de cette détection sont disponibles sur la figure C.2.

## C.3 Caractérisation des naevus

Une fois la détection des régions d'intérêt contenant des naevus effectuée, il est alors possible de les caractériser. Pour y parvenir, on décide d'appliquer virtuellement les critères ABCD (Asymétrie, Bordures, Couleur et Diamètre) utilisés par les médecins. Chacune de ces étapes va retourner un indice de confiance par rapport au critère, puis un indice global (moyenne des quatre indices) va être renvoyé à l'utilisateur. Si l'indice global se rapproche de 0%, alors le naevus semble dangereux, s'il se rapproche de 100%, le naevus semble bénin.

### C.3.1 Asymétrie

Afin de déterminer directement l'asymétrie du naevus, quelques étapes de prétraitement sont effectuées pour préparer l'analyse de la symétrie. Dans un premier temps, le contour du naevus est détecté (figure C.3.a). A partir de ce contour, il est possible de

déterminer l'ellipse et son centre qui s'adapte le mieux au contour détecté (figure C.3.b). Puis, en effectuant la rotation de ce contour, il est alors possible de séparer le naevus en quatre éléments afin d'analyser la symétrie (figure C.3.c).

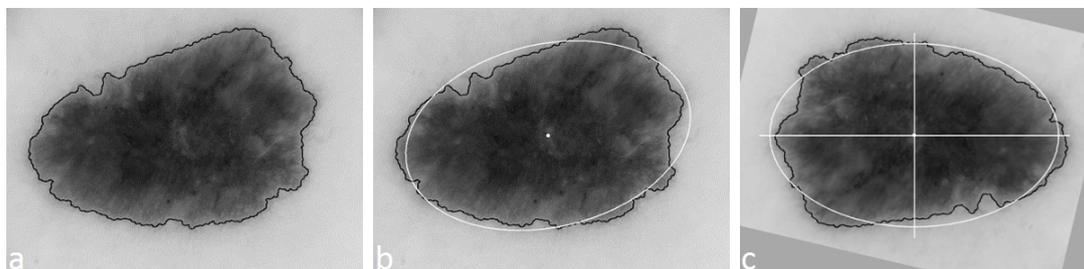


FIGURE C.3 – Asymétrie : Prétraitement, (a) Détection du contour du naevus, (b) Recherche du centre de gravité et de l'orientation du naevus, (c) Rotation et séparation du naevus

Le naevus est donc séparé en quatre éléments (figure C.4). Puis les premier et second moments de Hu [Hu62] sont appliqués à ces éléments. En effet, M. Hu montre qu'avec ses deux premiers moments, il est possible de savoir si deux éléments sont identiques. Il suffit alors de calculer ces deux moments pour les deux éléments, et s'ils égaux deux à deux, alors les deux éléments sont identiques. Et donc les deux éléments sont symétriques.

A partir de cela, un indice de confiance est retourné le critère de symétrie : plus il s'approche de 100% et plus il est symétrique.

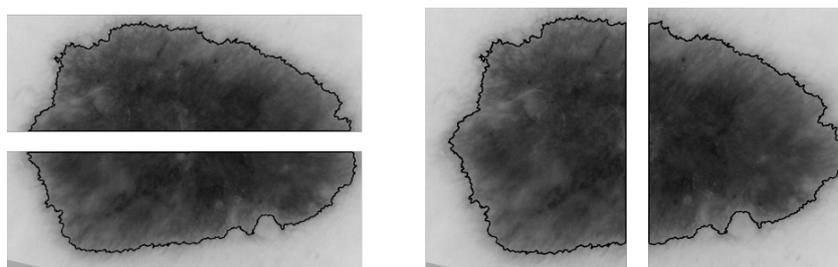


FIGURE C.4 – Asymétrie : Recherche de similarités entre deux demi naevus

### C.3.2 Bordures

Pour connaître la régularité des bordures, on va s'intéresser plus particulièrement au contour du naevus. On va donc calculer le périmètre du contour du naevus (figure C.5.a) et on détermine aussi le périmètre du contour convexe du naevus (figure C.5.b). On suppose donc ici que le contour convexe du naevus est le contour sans la moindre irrégularité, et comparé au contour réel, on est capable de traduire les irrégularités de bordures du naevus. Cela se fait en regardant la différence entre les deux périmètres de ces contours

(figure C.6), et plus la différence est élevée, plus les bordures du naevus sont irrégulières. Grâce à cela, on peut retourner un indice de confiance traduisant la régularité des bordures : proche de 0% si les bordures sont très irrégulières et 100% si elles sont régulières.

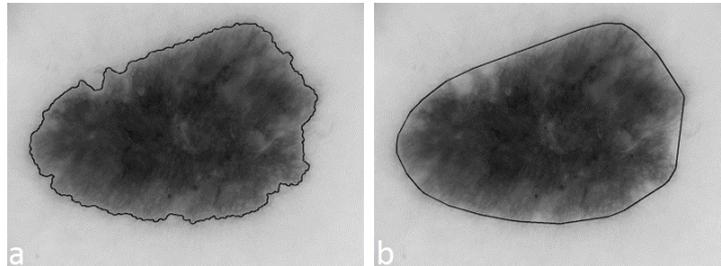


FIGURE C.5 – Bordures : Détermination de contours, (a) Contour au plus près du naevus, (b) Contour convexe du naevus

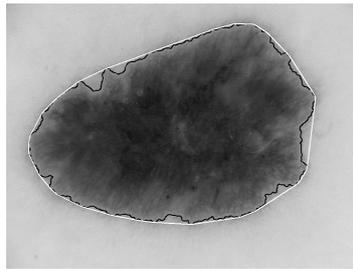


FIGURE C.6 – Bordures : Comparaison des longueurs des contours

### C.3.3 Couleur

Concernant la couleur du naevus, on procède de la façon suivante : on découpe l'image au plus près du naevus (figure C.7) puis on calcule la matrice de cooccurrence associée à cette image. La matrice de cooccurrence [HS<sup>+</sup>73] permet de représenter la distribution de valeurs de pixels cooccurrents dans une direction donnée.

A partir de la matrice de cooccurrence (figure C.8), il est possible de calculer l'homogénéité de notre image, et donc de la couleur du naevus. L'homogénéité est obtenue en appliquant la formule suivante :  $homogénéité = \sum_{i,j=0}^{255} \frac{P_{i,j}}{1+(i-j)^2}$  avec  $P_{i,j}$  qui correspond à la valeur contenue dans la matrice aux coordonnées  $(i, j)$ . Si la matrice est normalisée (qui est le cas ici), le résultat de l'homogénéité est compris entre 0 et 1, qui permet alors de renvoyer un indice compris entre 0% et 100%.

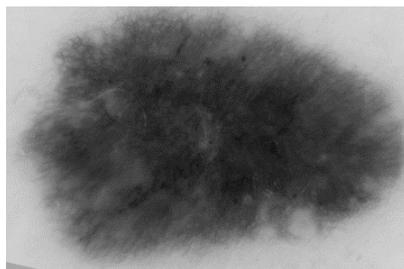


FIGURE C.7 – Couleur : Illustration d'un naevus recadré

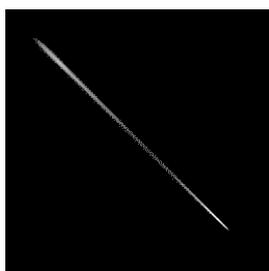


FIGURE C.8 – Couleur : Matrice de cooccurrence associée à la figure C.7

### C.3.4 Diamètre

Pour déterminer le diamètre réel du naevus, on va se baser sur la formule disponible figure C.9. La formule permet de déterminer la taille exacte de l'objet en se basant sur plusieurs paramètres : la distance entre le naevus et l'objectif photo (fixée par l'utilisateur), la taille du naevus sur l'image en pixel, la taille du capteur photo, la longueur de la focale de l'objectif et la taille de l'image. La plupart de ces informations étant disponibles dans les données EXIF de l'image. Une fois le diamètre du naevus connu, on le compare à la valeur de 6 millimètres. Si la valeur est supérieure à 6 millimètres, alors on retourne 0% comme indice de confiance, sinon, on retourne 100%.

$$\text{taille réelle}(mm) = \frac{\text{distance de l'objet}(mm) \times \text{taille de l'objet}(pixel) \times \text{taille du capteur}(mm)}{\text{longueur de focale}(mm) \times \text{taille de l'image}(pixel)}$$

FIGURE C.9 – Diamètre : Formule pour déterminer le diamètre réel de l'objet

## C.4 Performances de cette technique

On s'intéresse donc maintenant aux performances de la méthode et voir si elle est adaptable au problème de détection des naevus. Pour cela, on utilise quelques unes des métriques présentées précédemment qui sont le rappel et la précision. Concernant les données utilisées pour tester cette méthode, on utilise les 300 images mises de côté à cet effet.

| Métriques | Après détection | Après caractérisation |
|-----------|-----------------|-----------------------|
| Rappel    | 98,67%          | 98,57%                |
| Précision | 98,01%          | 68%                   |

TABLE C.2 – Performances de la méthode de détection et caractérisation des naevus

Les résultats de ce test sont disponibles dans le tableau C.2.

D'après les résultats, on est capable de détecter plus de 98% des naevus avec un très faible nombre de fausses détections (précision de 98%). A propos de la caractérisation, on constate qu'on détecte 98% des naevus, cependant on en classe correctement que les deux tiers.

## C.5 Conclusion

On a donc mis au point une méthode de détection et de caractérisation des naevus basée sur la méthode développée pour la détection des polypes. Cette méthode tire donc parti de tous les avantages de la méthode originale, c'est-à-dire une capacité à fonctionner en temps réel, avec de bonnes performances de détection et facilement embarquable. On peut très bien envisager de porter cette méthode sur des plateformes embarquables (comme la Raspberry Pi) ou développer une application smartphone dédiée. Cela montre à quel point notre méthode de détection des polypes peut être adaptée à d'autres problèmes d'ordre médicaux.





# Résumé

Ce manuscrit présente les travaux réalisés pour la mise au point d'un nouvel outil dédié à la détection temps réel assistée par ordinateur pour la détection des polypes colorectaux en vidéocoloscopie dans le cadre du dépistage précoce du cancer colorectal. Ce cancer reste aujourd'hui associé à un taux de mortalité important lorsque détecté trop tardivement. Un tel outil pourrait donc permettre d'améliorer la prise en charge.

Tout d'abord, une analyse de l'état de l'art permet de mettre en évidence les limitations des méthodes actuelles. Il est alors possible de déterminer le positionnement scientifique et technique à adopter pour créer un outil innovant dédié à cette tâche de détection.

Une première phase consacrée à l'analyse d'images fixes permet de confirmer ou d'infirmer les choix scientifiques possibles et ainsi développer une approche en accord avec les contraintes imposées par une utilisation clinique (temps de traitement, performances en particulier). Cette première méthode est ainsi capable de détecter les polypes en moins de 40 millisecondes par image, soit compatible avec le temps réel, tout en ayant de bonnes performances ( $F_1$  Score de 47.55%).

Logiquement, on s'intéresse ensuite à la détection des polypes dans les vidéos, qui permet de se rapprocher d'examens réels. En particulier, on propose plusieurs optimisations (comme la cohérence temporelle) permettant de tirer bénéfice de la vidéo. L'outil est alors capable d'alerter le médecin sur la présence de polypes dans 97.6% des cas (36 vidéos) en temps réel avec, en moyenne, seulement 23 millisecondes pour traiter une image.

Au final, cet outil, capable d'alerter le médecin sur la présence de polypes colorectaux en temps réel durant l'examen, pourrait donc s'intégrer dans la routine clinique de la vidéocoloscopie. Des essais cliniques à venir permettront d'identifier les possibles limitations à lever.

# Abstract

This manuscript presents the work carried out for the development of a new tool dedicated to real-time computer-assisted detection for the detection of colorectal polyps in videocoloscopy, in the context of early detection of colorectal cancer. This cancer remains today associated with a high mortality rate when detected too late. A such tool could help to improve medical care.

First of all, an analysis of the state of the art makes it possible to highlight the limitations of current methods. It is then possible to determine the scientific and technical positioning to adopt in order to create an innovative tool dedicated to this detection task.

A first phase dedicated to the analysis of still frames confirms or invalidates the possible scientific choices and thus develops an approach in accordance with the constraints imposed by clinical use (processing time, performance in particular). This first method is able to detect polyps in less than 40 milliseconds per image, which is compatible with real time, with good performance ( $F_1$  Score of 47.55%).

Logically, we are interested in the detection of polyps in videos, which allows to get closer to real exams. In particular, several optimizations (such as temporal coherence) are proposed to benefit from video. The tool is then able to alert the doctor about the presence of polyps in 97.6% of cases (on a set of 36 videos) in real time with only an average processing time of 23 milliseconds per frame.

In the end, this tool, which can alert the doctor of the presence of colorectal polyps in real time during the examination, could therefore be part of the clinical routine of videocoloscopy. Future clinical trials will identify possible limitations to be overcome.