



Smart Vision Chip pour l'exploration du côlon

Orlando Chuquimia

► To cite this version:

Orlando Chuquimia. Smart Vision Chip pour l'exploration du côlon. Imagerie médicale. Sorbonne Université, 2019. Français. NNT : 2019SORUS192 . tel-02494557v2

HAL Id: tel-02494557

<https://theses.hal.science/tel-02494557v2>

Submitted on 12 Feb 2021

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.



SORBONNE UNIVERSITÉ

École Doctorale Informatique, Télécommunications et Électronique
ED130

Laboratoire d'informatique de Paris 6 – Département System on Chip

Smart Vision Chip pour l'exploration du colon

Thèse de Doctorat

Par Orlando CHUQUIMIA

Dirigée par P^r Bertrand GRANADO et P^r Xavier DRAY

Présentée et soutenue publiquement le 26 novembre 2019

Devant un jury composé de :

Virginie FRESSE	Université Jean-Monnet-Saint-Étienne	Rapporteur
Emmanuel BOUTILLON	Université de Bretagne Sud	Rapporteur
Olivier ROMAIN	Université de Cergy-Pontoise	Examineur
David BOL	Université Catholique de Louvain	Examineur
Marie-Aude VITRANI	Sorbonne Université	Examineur
Bertrand GRANADO	Sorbonne Université	Co-Directeur
Xavier DRAY	Sorbonne Université	Co-Directeur
Andrea PINNA	Sorbonne Université	Co-Encadrant

DÉDICACE :
À la mémoire de ma mère[†]

Qui était le motif et m'a donné la force de ne jamais abandonner et de
lutter à tout moment.

Orlando Chuquimia

Remerciements

J'exprime ma profonde gratitude aux membres de jury d'avoir accepté de participer à l'évaluation de mes travaux de thèse, David Bol, Marie-Aude Vitrani, Olivier Romain, spécialement Emmanuel Boutillon et Virginie Fresse pour le temps consacré à l'examen de mon manuscrit de thèse.

Je remercie mes directeurs de thèse Bertrand Granado et Xavier Dray et mon encadrant Andrea Pinna pour m'avoir fait confiance et avoir cru en moi, pour leurs encadrements durant ces trois années de thèse, leur soutien moral et leur dévouement pour les relectures. J'ai appris la bonne méthodologie de recherche avec rigueur scientifique et approche formelle. J'ai également grandement apprécié leurs qualités humaines.

J'exprime ma profonde gratitude à Emmanuel Boutillon et Christophe Marsala qui ont formé le comité de suivi de ma thèse pour m'avoir accompagné dans chaque étape de ma thèse et m'avoir apporté leurs conseils et leurs visions externes.

Merci à tous mes collègues de l'équipe Syel présents : Julien Denoulet, Khalil Hachicha, Farouk Vallete, Olivier Tsiakaka et passés : Chen Chen, Alexandre Brière, Laurent Lambert, Imen Dhif et mes camarades Faten Sahel, Valentin Rebiere et Meriem Kemache pour les moments partagés ensemble comme l'auberge Espagnole ou les séminaires d'équipe.

Plus spécialement, je remercie également mes collègues et ex-collègues du LIP6 : Alisson Gongora, Abdelkrim Bessaad, Santiago Cortijo, Pedro Lusich, Weiqin Xu, Kashi Mishra et Leo Orozco. Vous avez élargi mes horizons scientifiques et culturels.

Enfin, je remercie avec tout mon cœur ma mère, mon père, mes frères et ma sœur. Les encouragements, l'amour et le soutien dont ils ont fait preuve ont énormément contribué à ma motivation.

Ce travail a été effectué dans le cadre d'une bourse attribuée pour l'Institut Universitaire d'Ingénierie en Santé (IUIS) au sein du Laboratoire d'Informatique de Paris 6 (LIP6) de Sorbonne Université à Paris en France.

THÈSE DE DOCTORAT DE
SORBONNE UNIVERSITÉ

Smart Vision Chip pour l'exploration de colon

Résumé :

Le **Cancer Colorectal** (**Cancer Colorectal (CCR)**) est le deuxième les plus fréquent en terme de mortalité avec un taux de 47% (880 792 décès en 2018). 96% des **CCR** débutent par l'apparition d'une tumeur (un **polype adénomateux**) dans la paroi du côlon.

Le **CCR** est un cancer qui a de bon pronostic, avec des guérisons dans plus 90% des cas, à condition qu'il soit détecté tôt. Il représente donc un **enjeu majeur de santé publique** où il est nécessaire de mettre en œuvre des politiques de **dépistage précoce** afin de réduire sa mortalité.

La **Capsule Endoscopique (CE)**, a été inventée par Paul Swain en 1990. Il s'agit d'une pilule intégrant une caméra et un système de radio communication, que le patient avale. Pendant cette procédure, la capsule transmet les images du tube digestif vers un boîtier externe via une communication par radiofréquence. Une fois toutes les images transmises, un médecin gastro-entérologue les télécharge pour réaliser une analyse visuelle et détecter les anomalies et les tumeurs. À l'aide de ce dispositif les médecins peuvent détecter des polypes, d'au moins 5 mm, avec une sensibilité et une spécificité respectivement de 68,8% et de 81,3%. Cette **CE** présente quelques limites et faiblesses liées à la résolution spatiale et temporelle des images, à son autonomie énergétique et au grand nombre d'images transmises que le gastro-entérologue doit analyser.

Nous avons étudié la conception d'un système embarqué contenant une chaîne de traitement capable de détecter des polypes et intégrée dans une capsule, créant un nouveau dispositif médical : une **capsule endoscopique intelligente**. Pour réaliser ce dispositif, nous avons pris en compte toutes les contraintes non fonctionnelles liées à l'intégration dans une **capsule endoscopique**. Ce dispositif pourrait ainsi être un nouvel outil de dépistage précoce des lésions précancéreuses colorectales : les **polypes**.

La première partie de ces travaux de thèse est dédiée au développement d'une chaîne de traitement pour la détection automatique des polypes pouvant être intégrée dans une capsule. La deuxième partie de ces travaux est dédiée à la conception d'un système embarqué intégrable dans la capsule et contenant la partie la plus critique de notre chaîne : la transformée de Hough.

Les résultats obtenus dans cette thèse sont à l'état de l'art et contribuent au développement de la nouvelle génération de capsule endoscopique. Ils font preuve de la capacité de calcul qu'il est possible d'intégrer efficacement dans un environnement très contraint (temps, surface et énergie).

Mots clefs : Détection automatique des polypes, capsule endoscopique intelligente, transformée d'Hough, arbres flous, forêts floues, vision artificielle et accélérateur matérielle.

PHD THESIS OF
SORBONNE UNIVERSITÉ

SMART VISION CHIP FOR COLON
EXPLORATION

Abstract:

CCR is the second highest cause of death by cancer worldwide with 880,792 deaths in 2018 and a mortality rate of 47.6%. 95% of CCR cases begin with the presence of a growth on the inner lining of the colon or the rectum, called a polyp. Multiple types of polyps exist; among them, adenoma polyps, which can degenerate into CCR. CCR is treatable in 90% of the cases if it is detected early enough. This is a real public health problem where it is necessary to implement early detection policies to reduce the mortality rate of this cancer.

The endoscopic capsule was invented by Paul Swain in 1990. It is a pill incorporating a camera and a radio communication system that the patient swallows and transmit images from the gastrointestinal tract through the body in a workstation. Once all images are transmitted, a gastroenterologist downloads them to perform a visual analysis and detect abnormalities and tumors. Using this device doctors can detect polyps, at least 5 mm, with sensitivity and specificity respectively of 68.8% and 81.3%. This endoscopic capsule presents some limitations and weaknesses related to the spatial and temporal resolution of images, its energy autonomy and the number of images transmitted to be analyzed by the gastroenterologist.

We studied the design of an embedded system containing a processing chain capable of detecting polyps to be integrated into an endoscopic capsule, creating a new medical device: an intelligent endoscopic capsule. To realize this device, we took into account all the non-functional constraints related to the integration into an endoscopic capsule. This device must be a new tool for early detection of precancerous colorectal lesions: polyps.

The first part of this thesis is dedicated to the development of a processing chain for the automatic detection of polyps that can be integrated into a capsule. The second part of this work is dedicated to the design of an embedded system to be integrated into the capsule and containing the most critical part of our chain: the Hough transform. The results obtained in this thesis are at the state of the art and contribute to the development of the new generation of the endoscopic capsule. They demonstrate the computational capacity that can be effectively integrated into a highly constrained environment (time, surface and energy).

Keywords: Automatic polyp detection, intelligent endoscopic capsule, Hough transform, fuzzy trees, fuzzy forest, artificial vision and hardware accelerator.

Table des matières

Résumé	iv
Abstract	v
Table des matières	v
Table des figures	viii
Liste des tableaux	x
Acronyms	xiii
I. INTRODUCTION ET CONTEXTE	1
Introduction générale	3
Description des travaux	6
Apports des travaux	8
Publications et communications	9
1 Contexte et problématique	11
1.1 Le Cancer Colorectal (CCR)	12
1.2 Dépistage de cancer colorectal	15
1.3 CE disponibles dans le marché :	22
1.4 Limites actuelles de la CE :	24
1.5 L’avenir de la CE	26
Synthèse et problématique	28
II. DÉTECTION AUTOMATIQUE DES POLYPES	30
Introduction	33
2 Contexte et état de l’art	35
2.1 Méthodes pour la reconnaissance automatique des polypes	37
2.2 Analyse	44

3	Classification des polypes par arbres flous	47
3.1	Extraction des ROI	49
3.2	Description des ROI	50
3.3	Apprentissage des arbres flous	52
3.4	Test de la classification des ROI avec les arbres flous	55
3.5	Classification des ROI avec une forêt d'arbres flous	58
3.6	Discussion	59
4	Extraction automatique des ROI	61
4.1	Experimentations	63
4.2	Extraction et suivi des ROI pour la détection des polypes	67
4.3	Discussion	72
5	Entrainement du classifieur en utilisant les ROI extraites par la chaîne de traitement	73
5.1	Expérimentations	74
5.2	Un approche de Boosting pour augmenter la performance de détection	77
5.3	Mesures au niveau polype et patient	81
	Synthèse et discussion	83
III.	MISE EN ŒUVRE MATÉRIELLE	85
	Introduction	87
	Profilage de la chaine de traitement sur un processeur embarqué	88
6	État de l'art : transformée de Hough et ses implémentations	92
6.1	Algorithme original de la transformée de Hough	93
6.2	Algorithme de la transformée d'Hough Unidimensionnelle (1DHT) . .	93
6.3	Synthèse et analyse	96
7	Mise en œuvre de l'algorithme 1DHT en temps réel	98
7.1	Accélération de calcul des centres candidates	98
7.2	Architecture pour la mise en œuvre de l'algorithme 1DHT modifié . .	103
7.3	Mise en œuvre de l'architecture du 1DHT IP sur un FPGA	109
7.4	Mise en œuvre de l'architecture du 1DHT IP dans un Système sur Puce	112
8	Optimisation de l'IP 1DHT en surface	118
8.1	Optimisation du calcul des centres candidats	118
8.2	Architecture proposée pour la mise en œuvre du calcul des coordon- nées en x	121
8.3	Intégration de l'IP 1DHT-optimisée sur FPGA	125
8.4	Intégration de notre IP 1DHT-optimisée dans une système sur puce .	125
8.5	Estimation de la consommation de l'IP optimisée	127
	Synthèse et discussion	129

IV. CONCLUSION	131
9 Contributions et perspectives	132
9.1 Contributions	133
9.2 Perspectives	134
V. ANNEXES	135
A Annexes	136
A.1 Descripteurs extraits à partir de la matrice de co-occurrence et de l'histogramme de luminosité	136
Bibliographie	141

Table des figures

1	Incidence et mortalité du cancer dans le monde.	4
2	Taux d'incidence du cancer colorectal dans le monde.	5
1.1	Les stades du Cancer Colorectal (CCR) , source [6]	13
1.2	Dépistage du CCR et taux de mortalité d'après [64].	15
1.3	Pourcentage de dépistage à KPNC (Source : [64])	18
1.4	Coloscopie (Source : National Cancer Institute)	18
1.5	Capsule endoscopique : PillCam	21
2.1	Arbre de décision	38
2.2	Machines à vecteurs de support	39
2.3	a) image originale, b) image de contours, c) accumulation des centres de courbure et d) segmentation de l'image.	41
2.4	a) image originale, b) image de contours, c) gradients d'intensité d) accumulation des gradients d'intensité.	41
2.5	Algorithme de Local Binary Pattern	42
2.6	Algorithme Uniform Binary Pattern	43
3.1	Arbre Flou	47
3.2	Analyse des arbres flous	49
3.3	a) image , b) masque binaire associé à l'image.	50
3.4	Extraction des Region Of Interest (ROI) et calcul des descripteurs.	51
3.5	Algorithme de construction d'un arbre flou.	52
3.6	a) Valeurs de l'attribut A_i , b) fonction d'appartenance de l'attribut A_i discrétisé	53
3.7	Résultats de classification pour 50 Arbres Flous.	57
3.8	Forêt Flou	58
3.9	Résultats de la classification avec la forêt de 50 arbres flous.	59
4.1	Diagramme d'une chaîne de traitement pour la reconnaissance des polypes.	61
4.2	Étape d'extraction des ROI	63
4.3	a) ROI extraites à la main, b) ROI extraites automatiquement.	65
4.4	Sensibilité de classification selon le pourcentage de polype contenu dans une ROI	66
4.5	Chaîne de traitement pour l'extraction et suivi des ROI pour la dé- tection automatique des polypes.	67
4.6	Agrégation des ROI	69

4.7	a) Correspondance des blocs avec un voisinage d'une distance de 2 pixels. b) Correspondance de blocs avec 2 voisinages d'une distance de 2 et de 4 pixels. c) ROI validé dans l'image I_{n-1} . d) 8 vecteurs candidats de mouvement. e) 8 blocs candidats déplacés par les vecteurs candidats de mouvement. f) Bloc ayant une somme des valeurs absolues , SAD_{ij} , minimale.	70
5.1	Méthode proposée pour la construction de la base de données des ROI.	74
5.2	Résultats de détection des trois forêts floues sur la base de test. . . .	75
5.3	Résultats de reconnaissance des trois forêts floues sur la base d'apprentissage.	76
5.4	Exemple a) d'un modèle qui a sur-appris, b) d'un modèle qui a sous-appris et un modèle de séparation quasi-idéal (ligne noire).	76
5.5	Résultats des performances de reconnaissance des forêts floues boostées sur la base d'apprentissage et sur la base de test.	79
5.6	Mesures au niveau polype et au niveau patient.	81
5.7	Architecture Zynq-7000 [9].	88
5.8	Carte de développement Zedboard (source : www.digilentinc.com). . .	89
5.9	Intégration dans la carte de développement Zedboard.	89
6.1	a) Calcul des x-cordonnées. b) Calcul des y-cordonnées. c) Calcul des rayons.	95
7.1	Validation de l'algorithme avec différentes valeurs du seuil f	101
7.2	Image synthétique sans bruit avec 10 cercles non superposés.	102
7.3	Image synthétique sans bruit avec 10 cercles superposés.	102
7.4	Image synthétique avec bruit avec 10 cercles non superposés.	102
7.5	Synoptique de l'architecture proposée 1DHT IP.	104
7.6	Architecture pour le calcul des coordonnées en x.	105
7.7	Architecture pour le calcul de rayon.	107
7.8	Architecture FPGA pour le module calcul des x-coordonnées . . .	110
7.9	Architecture FPGA pour le calcul de rayon.	111
7.10	Diagramme de l'IP 1DHT intégré dans un système Zynq.	113
7.11	Intégration de l'IP "AXI 1DHT" sur Zynq : pipeline au niveau système	114
8.1	a) image réel, b) image de contour et c) cercle détecté en vert.	121
8.2	Architecture optimisée en ressources pour le calcul des coordonnées en x.	122
8.3	a) Image de contour, histogramme H_x calculé par b) l'architecture x-coordonnées IP 1DHT et par c) la nouvelle architecture IP 1DHT optimisée. Dans les histogrammes l'axe x et y correspondent aux coordonnées en x et en y respectivement	124
8.4	Cercles détectés	126

Liste des tableaux

1.1	Sensibilité et spécificité de détection de polypes du test immunologique Fecal Immunochemical Test (FIT) en fonction de sa sensibilité à la concentration en hémoglobine	17
1.2	Fréquence des adénomes et des polypes non adénomateux durant deux examens par coloscopies.	19
1.3	Capsules endoscopiques disponibles sur le marché	22
1.4	Performances des CE	24
1.5	Piles bouton dans le marché	26
2.1	Détails de ASU-Mayo Clinic Colonoscopy Database [89].	36
2.2	Classification de Paris	37
2.3	Détails de la base d'apprentissage CVC-VideoClinicDB [25, 20].	37
2.4	État de l'art : méthodes de détection automatique des polypes	45
3.1	Opérateurs de norme \top et conorme \perp	56
4.1	Détails de partage de la base de données des images de coloscopie ASU-Mayo [89].	63
4.2	Performances de détection de polypes	64
4.3	Performances de détection de polypes	71
5.1	Profil des temps d'exécution de la chaîne de traitement pour des images Full High Definition, 1920x1080 (FHD)	90
6.1	Performances de l'état de l'art pour la transformée de Hough.	96
7.1	Comparaison entre notre algorithme et la référence de l'algorithme de la transformée de Hough de la librairie OpenCv.	103
7.2	Calcul des coordonnées en x et des coordonnées en y pour une image de taille $H * W$	108
7.3	Calcul des rayons pour une image de taille $H \times W$	109
7.4	Ressources des modules Calcul des x-coordonnées et Calcul des y-coordonnées pour une image de taille $H * W$	112
7.5	Calcul des rayons pour une image de taille $H \times W$	112
7.6	Caractéristiques de mise en œuvre de l'IP "AXI 1DHT IP" pour une image de taille $H \times W$	115
7.7	Mise en œuvre de l'IP "AXI 1DHT IP" sur le Zynq.	115
7.8	Notre méthode pour différentes tailles d'image.	116
7.9	Comparaison de cadence de notre méthode avec l'état de l'art.	116
7.10	Résultats de la mise en œuvre de l'IP 1DHT sur le Spartan 7 XC7S15.	117

7.11	Ressources nécessaires pour la mise en œuvre de l'IP 1DHT	117
8.1	Comparaison de notre algorithme optimisé en surface par rapport à la transformée de Hough de la librairie OpenCv	121
8.2	Ressources nécessaires pour le calcul des coordonnées en x et en y pour une image de taille $H * W$	123
8.3	Estimation des ressources nécessaires pour le calcul des coordonnées en x et en y.	125
8.4	Résultats d'implantation de l'IP 1DHT-optimisée sur le FPGA Xilinx Spartan 7 XC7S15.	125
8.5	Détails de la mise en œuvre de l'IP "AXI 1DHT-optimisée".	126
8.6	Autonomie énergétique estimée de l'IP "AXI 1DHT-optimisée" dans une CE	127
8.7	Mise en œuvre de notre IP 1DHT-optimisée dans une technologie CMOS Application Apecific Integrated Circuit (ASIC) de 65 nm. Avec une fréquence de 125 MHz.	128

Acronymes

1DHT	One Dimensional Hough Transform.
ADN	Acide Désoxyribo Nucléique.
AFLS	Adaptative Fuzzy Logic System.
AJCC	American Joint Committee on Cancer.
ANN	Artificial Neural Network.
ARM	Advanced RISC Machine.
ASIC	Application Apecific Integrated Circuit.
CCR	Cancer Colorectal.
CCR	Cancer Colorectal.
CE	Capsule Endoscopique.
CE	Capsule Endoscopique.
CNN	Convolutional Neural Network.
ECO	Efficient Convolution Operators.
FHD	Full High Definition, 1920x1080.
FIT	Fecal Immunochemical Test.
FN	Faux Négatif.
FP	Faux Positif.
FPGA	Field Programable Gates Array.
FPM	Generalized Fuzzy Partition using Mathematical Morphology.
GIANA	Gastrointestinal Image ANALysis.
HD	Haute Définition.
HOG	Histogram of Oriented Gradients.
HSV	Hue Saturation Value.
iCE	Capsule Endoscopique Intelligente.
KPNC	Kaiser Permanente Northern California.
LBP	Local Binary Pattern.

MICCAI	Medical Image Computing and Computer Assisted Intervention.
RGB	Red Green Blue.
ROI	Region Of Interest.
SD	Standard Définition.
SVSoC	Smart Vision System on Chip.
TNM	Tumor, Nodes, Metastasis.
VC	Virtual Colonoscopy.
VN	Vrai Négatif.
VP	Vrai Positif.

Première partie

INTRODUCTION ET CONTEXTE

Introduction générale	3
Description des travaux	6
Apports des travaux	8
Publications et communications	9
 1 Contexte et problématique	 11
1.1 Le Cancer Colorectal (CCR)	12
1.2 Dépistage de cancer colorectal	15
1.2.1 Test immunologique FIT	16
1.2.2 Coloscopie	17
1.2.3 Coloscopie Virtuelle (Virtual Colonoscopy (VC))	20
1.2.4 Capsule Endoscopique (CE)	20
1.3 CE disponibles dans le marché :	22
1.3.1 Medtronic : PillCam	22
1.3.2 Jinshan : OMOM	23
1.3.3 Olympus : EndoCapsule 10	23
1.3.4 CapsoVision : CapsoCam Plus	23
1.4 Limites actuelles de la CE :	24
1.5 L’avenir de la CE	26
 Synthèse et problématique	 28

Introduction générale

La recherche a besoin d'argent dans deux domaines prioritaires : le cancer et les missiles antimissiles. Pour les missiles antimissiles, il y a les impôts. Pour le cancer, on fait la quête.

Pierre Desproges
Artiste, Comique (1939 - 1988)

Le corps humain est un système biologique complexe composé de plusieurs systèmes biologiques tout aussi complexes appelés organes. Ceux-ci sont composés de différents tissus eux-mêmes composés des cellules qui grandissent, se divisent et se multiplient de manière ordonnée et contrôlée.

L' [Acide Désoxyribo Nucléique \(ADN\)](#) supporte l'information génétique des cellules. Quand les cellules sont endommagées ou qu'elles vieillissent, un processus appelé apoptose assure leur mort : on parle de mort cellulaire programmée.

Une des limites de l'apoptose est liée à sa défaillance en présence de certaines mutations génétiques. Les cellules endommagées dans ce cas grandissent sans contrôle générant des erreurs dans un système par ailleurs d'une grande perfection.

La multiplication incontrôlée de ces cellules entraîne une augmentation du volume qui est appelée [tumeur](#). Ces tumeurs peuvent être strictement bénignes, sans risque d'essaimage à distance. Elles peuvent être bénignes initialement, mais avec un potentiel dégénératif malin. Ou elles peuvent être malignes avec un risque d'essaimage vers des organes à distance, par voie sanguine ou lymphatique notamment. Localement, dans l'organe initialement concerné par la tumeur, il peut y avoir un dysfonctionnement comme cela peut n'être responsable d'aucun symptôme.

Au stade de la malignité, ou [cancer](#), les cellules tumorales de lésion « mère », ou « primitive », ont acquis la capacité de se propager vers d'autres tissus, en passant dans le sang ou les vaisseaux lymphatiques, et cette propagation peut former de nouvelles tumeurs « filles », ou « secondaires », à distance, appelées métastases. La nature tissulaire de ces métastases est souvent très similaire à celle de la lésion primitive. Les traitements des cancers varient souvent selon la nature de l'organe siège du cancer primitif, et de l'existence de métastases.

Aujourd'hui le cancer est une des principales causes de mortalité au monde.

Selon l'agence internationale de recherche du cancer¹ [12, 28], le **cancer colorectal (CCR)** est le troisième cancer le plus fréquent et le deuxième le plus mortel avec un taux de décès de 47%, comme nous pouvons le voir dans la Figure 1. Quelques 880 792 décès, sur 1 849 518 cas de **CCR**, ont été enregistrés en 2018 parmi lesquels 1 356 151 nouveaux cas.

En France il y a un décès par **CCR** toutes les 30 minutes en moyenne [12, 28]. Ce cancer est plus fréquent chez les hommes (55%), mais plus mortel pour les femmes (48%). Le risque augmente à partir de l'âge de 50 ans. L'âge moyen au moment du diagnostic est de 70 ans. Ce cancer est plus fréquent dans les pays développés, comme nous pouvons le constater dans la Figure 2.

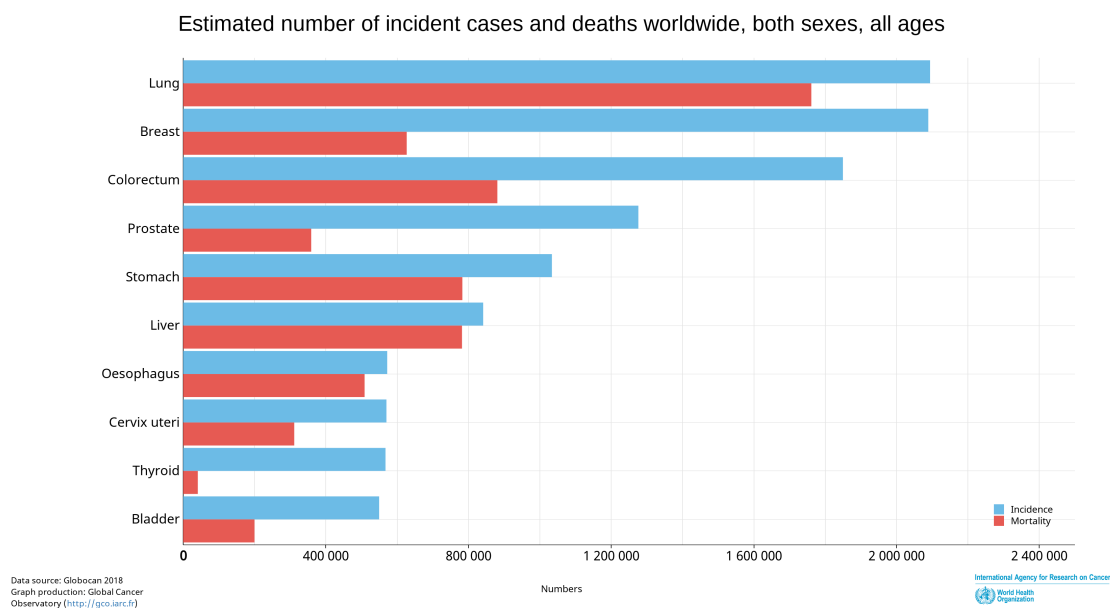


FIGURE 1 : Incidence et mortalité du cancer dans le monde.

96% des **CCR** commencent par l'apparition d'une tumeur dans la paroi du côlon ou du rectum. Cette tumeur est appelée **polype adénomateux ou adénome**. Le **CCR** peut être soigné dans 90% des cas à condition qu'il soit détecté dans les étapes précoces du développement du polype.

Le **CCR** est un **problème majeur de santé publique** pour lequel des politiques de **dépistage précoce** avec un haut niveau d'acceptation des tests de **dépistage** sont souhaitables.

Les travaux de doctorat décrits dans ce manuscrit concerne l'étude de la conception d'un système embarqué contenant une chaîne de traitement capable de détecter des polypes et pouvant être intégrée dans un nouveau dispositif médical, une

¹agence en charge de la collecte, de l'analyse, de l'interprétation et de la dissémination des indicateurs statistiques sur l'incidence, la prévalence et la mortalité de tous les types de cancer au monde

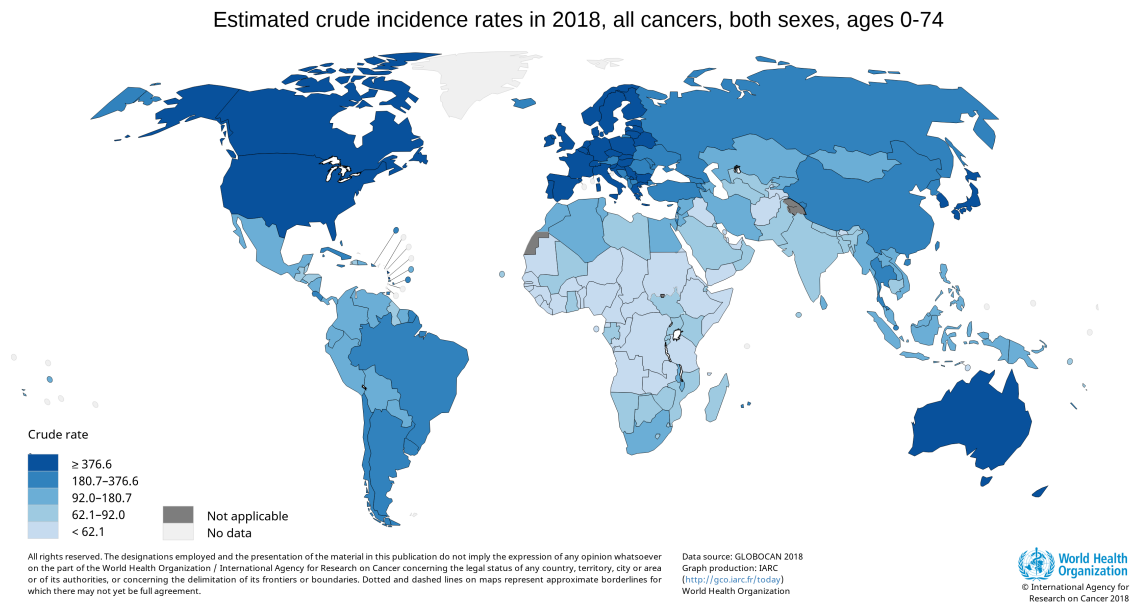


FIGURE 2 : Taux d'incidence du cancer colorectal dans le monde.

Capsule Endoscopique Intelligente (iCE). La problématique de ces travaux concerne la conception d'un système embarqué possédant de fortes contraintes non fonctionnelles liées à l'intégration dans une capsule endoscopique. Ce système pourrait permettre le diagnostic précoce des **polypes** précancéreux colorectaux.

Description des travaux

Ce manuscrit est organisé en quatre parties, **Introduction et contexte, détection automatique des polypes, mise en œuvre matérielle et conclusions et perspectives**, et 9 chapitres.

- Dans la Partie I., nous établissons notre problématique et introduisons notre proposition de réaliser un système embarqué intégrable dans une capsule vidéo-endoscopique réalisant un nouveau dispositif médical, une capsule endoscopique intelligente, pour le diagnostic précoce de lésions précancéreuses colorectales : les polypes.
- La Partie II. relate le développement d'une chaîne de traitement pour la détection automatique des polypes en prenant en compte les contraintes d'intégration dans une capsule. Cette partie contient les chapitres suivants :
 - dans le Chapitre 2 nous présentons un état de l'art des méthodes de reconnaissance automatique des polypes ;
 - dans le Chapitre 3 nous décrivons l'étude de la reconnaissance des polypes à l'aide d'un classifieur utilisant une forêt d'arbres flous ;
 - dans le Chapitre 4 nous présentons une chaîne de traitement pour réaliser une détection automatique des polypes utilisant les arbres flous ;
 - dans le Chapitre 5 nous expliquons les optimisations que nous avons réalisées sur notre chaîne de traitement pour augmenter ses performances de détection. Pour cela nous avons utilisé l'analyse que nous avons faite de l'expertise des gastro-entérologues.
- La Partie III. présente la conception d'une architecture numérique pour l'exécution en temps réel de la partie la plus critique de notre chaîne : la transformée de Hough. Cette partie est constituée des chapitres suivants :
 1. dans le Chapitre 6 nous présentons un état de l'art des mises en œuvre temps réel de la transformée de Hough dans un [Field Programable Gates Array \(FPGA\)](#) ;
 2. dans le Chapitre 7 nous montrons notre étude de la mise en œuvre de la transformée de Hough en temps réel. Cette étude a donné lieu à la réalisation et l'optimisation d'un nouvel algorithme pour le calcul de la transformée de Hough afin de satisfaire les contraintes de temps réel liées à la [iCE](#). A l'aide de cet algorithme, nous avons conçu une architecture matérielle pour le calcul de la transformée de Hough capable de traiter

des images hautes définitions en temps réel et l'avons mis en œuvre sur un système sur puce ;

3. dans le Chapitre 8 nous décrivons la mise en œuvre de la transformée de Hough en temps réel et dans la surface transversale de la capsule endoscopique ($8 \times 8 \text{ mm}^2$). Cela nous a amené à concevoir une nouvelle architecture matérielle optimisée pour le calcul de la transformée de Hough capable de traiter des images hautes définitions en temps réel et intégrable dans la surface d'une capsule endoscopique.
- Dans la Partie IV., le chapitre 9 vient conclure ces travaux de thèse et présente quelques perspectives pour faire aboutir la réalisation et l'utilisation de ce nouveau dispositif médical.

Apports des travaux

Par rapport à l'état de l'art, à ce jour, ces travaux de thèse présentent les apports suivants :

1. la mise en œuvre d'une forêt d'arbres flous comme classifieur dans un chaîne de traitement pour la reconnaissance des polypes ;
2. l'introduction d'une nouvelle méthode pour la détection automatique des polypes basée sur l'expertise des gastroentérologues et qui prend en compte l'information spatiotemporelle ;
3. la mise en œuvre matérielle de notre chaîne de traitement sur un système sur puce ;
4. la réalisation et l'optimisation d'un algorithme pour le calcul de la transformée de Hough en temps réel ;
5. la conception d'une architecture matérielle pour le calcul de la transformée de Hough capable de traiter des images [Haute Définition \(HD\)](#) en temps réel et sa mise en œuvre sur un système sur puce ;
6. la conception d'une architecture matérielle optimisée pour le calcul de la transformée de Hough capable de traiter des images [HD](#) en temps réel et intégrable dans la surface d'une capsule endoscopique et sa mise en œuvre sur un système sur puce.

Publications et communications

Mes travaux de thèse ont donné lieu à un certain nombre de publications et présentations.

Cinq conférences internationales

- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Polyps recognition using fuzzy trees. In Biomedical & Health Informatics (BHI 2017) IEEE EMBS International Conference on, pp. 9–12, IEEE, 2017.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Fpga-based real time embedded hough transform architecture for circles detection. In Conference on Design and Architectures for Signal and Image Processing (DASIP 2018), IEEE International Conference, pp. 31–36, Oct. 2018.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Polyp follow-Up in an Intelligent Wireless Capsule Endoscopy. In Conference on In Conference on Biomedical Circuits and Systems (BIOCAS 2019), IEEE International Conference, pp. 31–36, Oct. 2019.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : A Real Time Hough Transform Architecture Useable inside a WCE. In Conference on In Conference on Biomedical Circuits and Systems (BIOCAS 2019), IEEE International Conference, pp. 31–36, Oct. 2019.
- T. Garbay, O. Chuquimia, A. Pinna, H. Sahbi, X. Dray and B. Granado : Distilling the knowledge in CNN for WCE screening tool. In Conference on Design and Architectures for Signal and Image Processing (DASIP 2019), IEEE International Conference, Oct 2019.

Deux challenges internationaux :

- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration. Sub-Challenge in Gastrointestinal Image Analysis GIANA at the 20th International Conference on Medical Image Computing and Computer Assisted Intervention, September 10th to 14th, Quebec City, Quebec, Canada, 2017.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration. Sub-Challenge in Gastrointestinal Image Analysis GIANA at the 21st International Conference on Medical Image Computing and Computer Assisted Intervention, September 16-20, Granada , Spain, 2018.

Sept conférences et colloques nationaux :

- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration. 1er Colloque Fédération d'Electronique de l'UPMC et de l'ISEP, novembre 2016, Paris, France.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration. Journée Embarquée pour le Médical : GDR ISIS, juin 2017, Lyon, France.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Polyps recognition using fuzzy trees. 12ème Colloque du GDR SoC/SiP 2017, juin 2017, Bordeaux, France.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration. 13ème Colloque du GDR SoC/SiP 2018, juin 2018, Paris, France.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip for Colon exploration to realize a screening tool, Ecole d'hiver Francophone sur les Technologies de Conception des Systèmes embarqués Hétérogènes, janvier 2019, Louvain-la-Neuve, Belgique.
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Study to integrate CNN inside a WCE to realize a screening tool, 7ème édition du colloque JETSAN 2019, May 2019, Paris, France. **(Obtention du prix du meilleur poster et de la meilleure présentation orale)**
- O. Chuquimia, A. Pinna, X. Dray, and B. Granado : Smart Vision Chip pour l'exploration du colon, Journée Challenges and trends in smart instrumentation and measurement organisée par le chapitre Frontière et interface Cyberphysiques IEEE I&M en association avec Le GDR SoC2, Oct 2019, Paris, France.

Chapitre 1

Contexte et problématique

You can't connect the dots looking forward; you can only connect them looking backwards. So you have to trust that the dots will somehow connect in your future. You have to trust in something – your gut, destiny, life, karma, whatever. Because believing that the dots will connect down the road will give you the confidence to follow your heart even when it leads you off the well worn path; and that will make all the difference.

Steve Jobs

Stanford commencement speech, June 2005

1.1 Le Cancer Colorectal ([CCR](#))

Le cancer colorectal apparait dans le côlon et le rectum. C'est une maladie qui se développe initialement à partir d'une tumeur bénigne, mais à potentiel malin, sur une période d'environ de 10 à 20 ans, sans provoquer de symptôme particulier, restant longtemps imperceptible. Certains des signes révélateurs d'un [CCR](#) sont les troubles persistants ou soudain du transit intestinal et les saignements.

Il y existe deux principaux types de polypes colorectaux :

- **le polype hyperplasique et inflammatoire** : ce type de polype est très fréquent et en général de petite taille et strictement bénin.
- **le polypes adénomateux ou adénome** : initialement bénin, son histoire naturelle est la dégénérescence en cancer (dans 96% des cas)

La détermination de la nature d'un polype est réalisée après son prélèvement qui est idéalement complet, sinon par un échantillonnage par biopsies. L'analyse visuelle, souvent textuelle, du tissu prélevé, au microscope, appelée analyse anatomopathologique ou histologique permet de réaliser l'analyse de l'échantillon. D'autres indicateurs comme la taille ou le nombre de polypes peuvent aussi être pris en compte.

Trois critères principaux sont pris en compte lors de l'analyse histologique pour juger du stade évolutif, c'est-à-dire le potentiel dégénératif, ou dégénéré, d'un polype adénomateux :

1. la taille et la profondeur d'envahissement du polype dans la paroi colique,
2. la présence et le nombre de ganglions lymphatiques atteints,
3. la présence de métastases.

Ces trois critères permettent de définir le stade de l'adénome ou du cancer selon la classification Tumor, Nodes, Metastasis ([Tumor, Nodes, Metastasis \(TNM\)](#))[6]
¹ ². Cette classification a été proposée par l'American Joint Committee on Cancer ([American Joint Committee on Cancer \(AJCC\)](#)). Cinq stades, visibles sur la figure 1.1, sont reconnus :

- **Stade 0** : l'adénome est in situ, ce qui signifie que la tumeur est très superficielle.
- **Stade I** : l'adénome envahit la couche sous-muqueuse ou la couche musculaire (muscleuse) de la paroi du côlon ou du rectum. Le pourcentage de mortalité à ce stade est de 6,1% [61].

¹<https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-colon>

²<https://www.cancer.org/es>

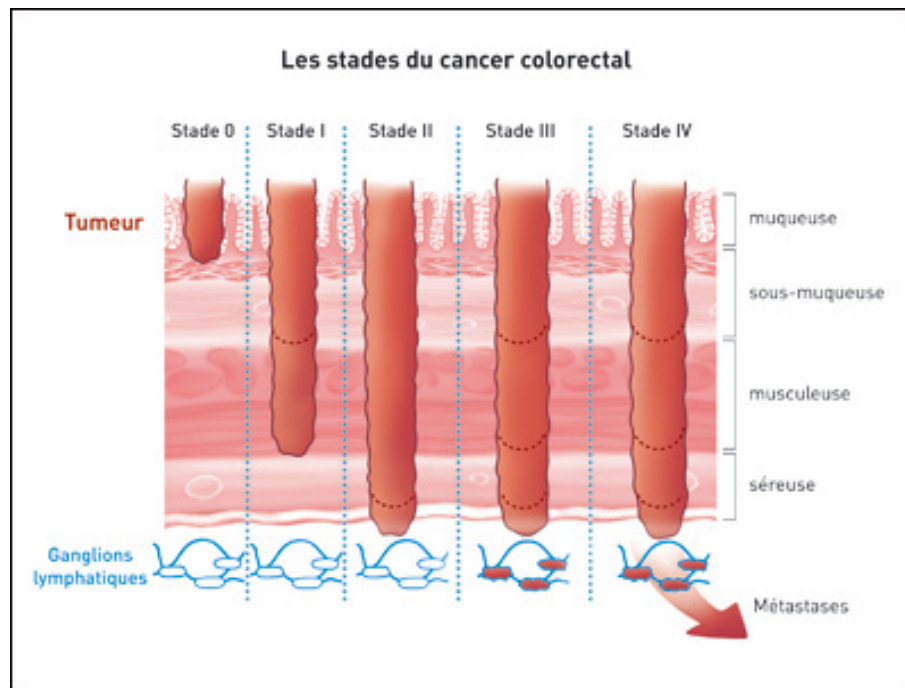


FIGURE 1.1 : Les stades du CCR, source [6]

- **Stade II :** les cellules cancéreuses ont traversé plusieurs couches de la paroi du côlon ou du rectum, mais aucun ganglion n'est atteint. Le pourcentage de mortalité à ce stade est de 17,9% [61].
- **Stade III :** les cellules cancéreuses ont envahi les ganglions lymphatiques proches de polypes. Le taux de mortalité à ce stade est de 38,5% [61].
- **Stade IV :** le cancer s'est propagé au-delà du côlon ou du rectum, dans la cavité abdominales, c'est un envahissement intrapéritonéal, et/ou vers des organes à distance qui sont généralement le foie ou les poumons. Le taux de mortalité à ce stade est de 92% [61].

Au stade de polype adénomateux encore bénin, le diagnostic puis le traitement par résection du polype par des méthodes peu agressives, comme la coloscopie, sont une prévention très efficace de toute évolution maligne.

Parvenue au stade de CCR, l'envahissement de la paroi intestinale par la tumeur peut être responsable de troubles du transit et de saignements, mais peut également rester peu ou pas symptomatique. Cette situation, lorsqu'elle a été identifiée assez tôt, relève souvent d'un traitement chirurgical, parfois doublé d'une chimiothérapie anticancéreuse. Au prix néanmoins lourd du traitement, le pronostic reste globalement bon.

Au stade de cancer métastatique, le tableau clinique est plus complexe. Il n'est pas rare que la maladie ne soit révélée qu'à ce stade très avancé, et dont le pronostic

est sombre.

La reconnaissance précoce des patients porteurs d'adénome est donc un enjeu majeur de santé publique, puisque la maladie est :

- fréquente ;
- bénigne et de bon pronostic si reconnue à un stade précoce ;
- maligne et de mauvais pronostic si reconnue à un stade tardif.

Cet enjeu de santé publique fait l'objet de politiques de dépistage (en population) et de diagnostic (à l'échelle individuelle).

La mise en place et l'efficacité de ces politiques de santé publique de dépistage est liée à trois principaux facteurs de la méthode de dépistage :

- Elle doit être peu coûteuse pour être déployée à grande échelle ;
- Elle doit être bien acceptée par la population cible : il est préféré des méthodes peu ou pas invasives, faciles à mettre en œuvre à l'échelle individuelle ;
- Elle doit être très spécifique c'est-à-dire ne pas identifier à tort comme indemne des patients pourtant porteurs d'adénome à risque d'évolution maligne. Cette spécificité peut être obtenue au détriment de la sensibilité et donc amener à tort à un examen plus poussé des patients qui seraient indemnes d'adénome.

Cette démarche de dépistage ne concerne que des patients sans symptôme. En cas de symptômes comme la modification récente du transit, traces de sang dans les selles ou une anémie, la démarche est une approche diagnostique, très différente.

Actuellement, en France et dans de nombreux pays occidentaux, il y a trois niveaux de risque d'adénome à potentiel malin et donc de CCR :

- risque moyen : les hommes et femmes de 50 à 74 ans, sans histoire personnelle ou familiale. Ce profil correspond à environ 80% des patients chez lesquels un CCR sera finalement diagnostiqué ;
- risque élevé : sujets ayant une histoire familiale d'adénome de grande taille ou de cancer colorectal avant 65 ans. Ce profil correspond à environ 15% des patients chez lesquels un CCR sera finalement diagnostiqué ;
- risque très élevé : patients ayant une prédisposition génétique au CCR. Parmi ces prédispositions il y a la polypose adénomateuse familiale et le syndrome de Lynch notamment, soit 3 à 5% des patients chez lesquels un CCR sera finalement diagnostiqué. Il y a aussi une maladie inflammatoire chronique de l'intestin comme la maladie de Crohn ou la rectocolite hémorragique, dans leurs formes anciennes et étendues.

Notons qu'il n'existe pas de catégorie à « risque faible ». Pour la population à risque moyen, le dépistage des adénomes colorectaux repose sur la recherche de sang dans les selles. Si ce test est positif, l'examen suivant est une coloscopie. Pour les populations à risque élevé et très élevé, le dépistage des adénomes colorectaux repose d'emblée sur la coloscopie, avec divers degrés de sophistication pour les patients à risque très élevé.

Dans la prochaine section, nous présentons ces tests, leurs avantages et leurs inconvénients.

1.2 Dépistage de cancer colorectal

Il est démontré que les politiques de dépistage précoce du CCR diminuent efficacement le taux de mortalité. Pour être plus précis, l'étude réalisée à Kaiser Permanente Northern California (KPNC) en Californie [64] montre que le pourcentage de dépistage est inversement proportionnel au taux de mortalité du CCR, voir la figure 1.2.

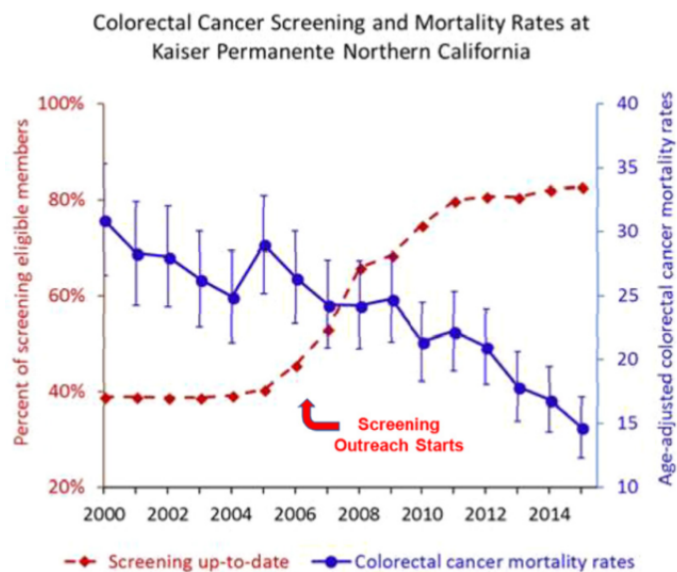


FIGURE 1.2 : Dépistage du CCR et taux de mortalité d'après [64].

Chaque pays adopte un programme de dépistage du CCR. En France, le programme de dépistage, généralisé depuis 2009, s'adresse tous les deux ans aux femmes et aux hommes de 50 à 74 ans sans histoire familiale ni antécédents personnels de CCR ou d'adénome et ne présentant pas de symptôme évocateur, population dite "risque moyen". Il est organisé en deux étapes. La première étape consiste en un **test immunologique** de recherche de saignement occulte dans les selles, test FIT.

Si le test est positif, dans 4 à 10% [35] des cas, un deuxième test plus poussé est prévu : la **coloscopie**. L'examen se déroule en principe sous anesthésie générale.

La coloscopie est l'examen de référence qui est pratiquée par un médecin gastroentérologue. Elle permet d'identifier la cause du saignement repéré par le test immunologique mais aussi, le cas échéant, de retirer d'éventuels polypes avant qu'ils ne se transforment en cancers.

Il existe d'autres méthodes de recherche d'adénomes comme la **Coloscopie Virtuelle (VC)** et la **Capsule Endoscopique (CE)**.

Les cas pour lesquels le gastroentérologue décide d'utiliser une méthode alternative à la coloscopie sont :

- en cas d'échec technique de la coloscopie, dans moins de 5% des cas, lorsque l'anatomie du côlon est complexe, le médecin ne parvient pas à faire un examen complet du côlon ;
- quand l'anesthésiste identifie des contre-indications pour le patient ;
- quand le patient n'accepte pas de se soumettre à une exploration coloscopique.

Les performances des tests de dépistages sont mesurées en terme de sensibilité, taux de vrais positifs, c'est à dire la capacité à détecter correctement un polype et en terme de spécificité, taux de vrais négatifs c'est à dire la capacité à ne pas se tromper. La sensibilité et la spécificité sont décrites par les équations suivantes :

$$Sensibilité = VP / (VP + FN) \quad (1.1)$$

$$Spécificité = VN / (VN + FP) \quad (1.2)$$

Où :

- VP : le taux de vrais positifs, c'est à dire le taux de tests déclarés positifs et qui sont positifs
- FP : le taux de faux négatifs, c'est à dire le taux de tests déclarés positifs et qui sont négatifs
- VN : le taux de vrais négatifs, c'est à dire le taux de tests déclarés négatifs et qui sont négatifs
- FN : le taux de faux négatifs, c'est à dire le taux de tests déclarés négatifs et qui sont positifs

1.2.1 Test immunologique FIT

Ce test consiste à détecter du sang dans les selles, résultant des micro-saignements dus à la possible présence d'un ou plusieurs adénomes.

Ce test est plus ou moins sensible selon la concentration d'hémoglobine qui se mesure en nano-grammes d'hémoglobine par milli-litre $\frac{ngHb}{mL}$ ou en micro-grammes d'hémoglobine par gramme de cellules $\frac{\mu gHb}{g}$.

TABLE 1.1 : Sensibilité et spécificité de détection de polypes du test immunologique **FIT** en fonction de sa sensibilité à la concentration en hémoglobine

Année, [étude]	Test de référence	Patients	Concentration HB ng/mL[ug/g]	adénome			CCR		
				n	Sens	Spec	n	Sens	Spec
2012, [35]	Coloscopie	1256	50 [10]	119	38%	93%	8	88%	91%
2012, [35]	Coloscopie	1256	75 [15]	119	33%	96%	-	-	-
2012, [35]	Coloscopie	1256	100 [20]	119	31%	97%	8	75%	95%
2013, [29]	Coloscopie	2235	- [24,5]	207	23%	97%	15	60%	95%
2013, [29]	Coloscopie	2235	- [7,95]	207	20%	97%	15	53%	95%
2013, [29]	Coloscopie	2235	- [6,1]	207	26%	97%	15	73%	96%
2014, [65]	Coloscopie	65	50 [50]	65	38%	88%	-	-	-
2014, [65]	Coloscopie	529	50 [6]	529	24%	85%	-	-	-
2014, [65]	Coloscopie	346	50 [10]	346	4%	97%	-	-	-
2014, [65]	Coloscopie	86	50 [50]	86	17%	89%	-	-	-
2015, [40]	Coloscopie	277	50 [50]	20	40%	-	3	75	90

Il existe plusieurs publications scientifiques qui présentent des résultats pour estimer la spécificité du test **FIT** sur une population en prenant comme test de référence la coloscopie. Certains de ces travaux sont listés dans le tableau 1.1. À la lecture de ce tableau, nous pouvons voir que la concentration d'hémoglobine a un impact direct sur la capacité de dépistage. Nous remarquons que le test **FIT** est très efficace pour dépister un **CCR** déclaré avec une spécificité de 90 à 96% et une sensibilité de 88%. Pour la détection d'un adénome ce test a une spécificité de 85 à 97% avec une sensibilité d'au mieux 40%. Dans ce second cas de figure, la répétition du test tous les 2 ans permet d'augmenter la sensibilité et d'identifier à temps la plupart des patients à risque.

Depuis son apparition en 2006, cette méthode a montré une augmentation du taux de dépistage du **CCR** par rapport aux autres méthodes. Cela a permis de passer de 39% de détection en 2000 à 82% en 2015, ceci notamment grâce à son haut niveau d'acceptation, voir la figure 1.3.

1.2.2 Coloscopie

La coloscopie est l'examen servant à visualiser l'intérieur du côlon : rectum, côlon sigmoïde, transverse, ascendant et descendant. Elle s'effectue à l'aide d'un coloscope pour chercher les polypes et notamment les adénomes, mais aussi toutes les autres anomalies. Le coloscope est constitué d'une source de lumière au Xénon et d'une caméra vidéo à haute résolution. Il est connecté à un écran qui permet au médecin

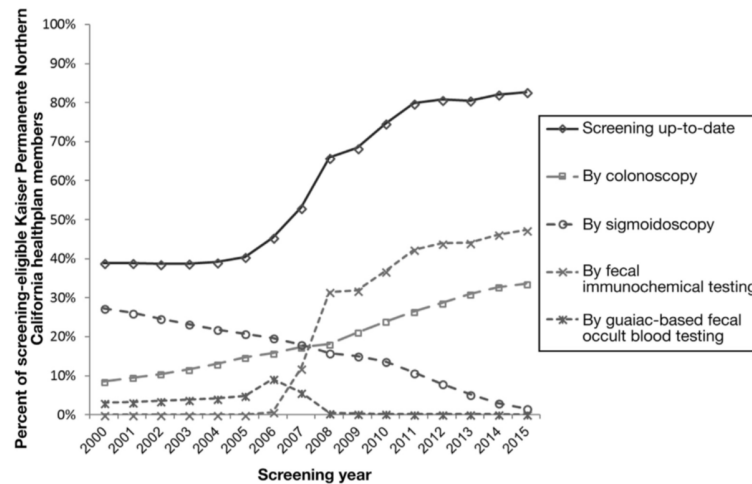


FIGURE 1.3 : Pourcentage de dépistage à KPNC (Source : [64])

de voir et d'enregistrer les images du côlon en temps réel. Le coloscope est constitué aussi d'un, ou plusieurs, canaux opérateurs, qui permettent d'insuffler du gaz (air ou dioxyde de carbone), de nettoyer le côlon et d'aspirer l'eau ou toute matière fécale restante. Ce canal permet aussi d'insérer des instruments pour réaliser des prélèvements ou des traitements, et notamment l'ablation des adénomes (voir la figure 1.4). La coloscopie est le test de référence utilisé par les gastroentérologues pour diagnostiquer et traiter les adénomes.

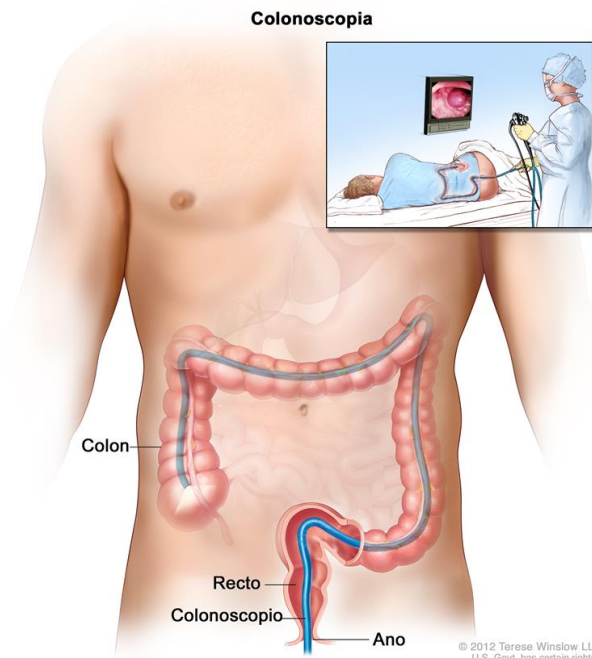


FIGURE 1.4 : Coloscopie (Source : National Cancer Institute)

Lors d'une coloscopie, le gastroentérologue trouve un polype dans 30-40% des

cas et un CCR dans 8% de cas [14].

Dans [76], les auteurs ont montré que la variabilité inter et intra médecins peut monter jusqu'à 26% dans la tâche de détection visuelle des adénomes et notamment les plus petits ³.

TABLE 1.2 : Fréquence des adénomes et des polypes non adénomateux durant deux examens par coloscopies.

Type	1er examen	2ème examen	Taux échec	Sensibilité
Adénomes < 5mm	217	81	27%	73%
Adénomes < 6-9mm	42	6	13%	87%
Adénomes > 10mm	30	2	6%	94%
Polypes non adénomateux	230	89	28%	72%
Total	519	178	34%	66%

Comment évaluer le taux de détection d'un polype en prenant en compte cette variabilité ?

Pour répondre à cette question, dans [77], les auteurs ont fait deux coloscopies consécutives le même jour sur 183 patients, les résultats sont listés dans la table 1.2. Ces résultats montrent une sensibilité élevée de 94% pour des adénomes les plus grands ayant une hauteur supérieure à 1cm, les plus à risque de cancer, mais une faible sensibilité de 73% pour des adénomes plus petits avec une hauteur de 5mm ou moins, les moins à risque.

Cette expérience a montré que les gastroentérologues ont des difficultés pour détecter les petits polypes.

Est-ce que ceci est lié en partie à la résolution de l'image ? Dans une autre étude [36], les auteurs ont montré que la haute-définition des images permet d'obtenir un meilleur taux de détection des adénomes, de 28,8% en HD à 24,3% en Standard Définition (SD), et des polypes, de 42,2% en HD à 37,8% en SD.

Malgré son taux de détection élevé la coloscopie est un examen **invasif**, inconfortable, et qui peut, dans de rares cas, amener des risques pour le patient tels que l'hémorragie et l'infection. Le principal avantage de la coloscopie, par rapport aux autres méthodes de dépistage, est la possibilité qu'elle offre de retirer les adénomes dans le même temps que leur détection, et de mettre le patient à l'abri de leur dégénérescence. Ceci est vrai à l'échelle de la population, ou du diagnostic à l'échelle individuelle. L'ablation d'adénome pendant la coloscopie revient à une opération par les voies naturelles et augmente son risque de complication telle que l'hémorragie ou la perforation, mais le bénéfice attendu qu'est la prévention du cancer contre-balance largement ces risques de complications.

³La variabilité inter médecins est la variabilité entre différents généralistes et la variabilité intra médecins est la variabilité de prescription au sein de la clientèle d'un même médecin.

Au total, il est démontré que la séquence test [FIT](#) puis coloscopie si le test [FIT](#) est positif diminue significativement le nombre et la mortalité du [CCR](#). Cette stratégie de dépistage diminue davantage la mortalité par [CCR](#) que les progrès réunis de la chirurgie, de la chimiothérapie et de la radiothérapie.

La principale limite de cette stratégie est qu'elle est très insuffisamment adoptée par la population : seuls 30% des patients chez lesquels un test [FIT](#) est indiqué s'y soumettent.

1.2.3 Coloscopie Virtuelle ([VC](#))

La coloscopie virtuelle, où colonographie, est un scanner à partir de laquelle une reconstruction du côlon en 3D est possible.

Afin d'effectuer cet examen, un agent de contraste est injecté au patient ainsi qu'un antispasmodique. Puis une sonde est introduite 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 à 30 secondes. L'examen dure environ 30 minutes.

Pour évaluer la sensibilité et la spécificité de détection de polypes de la coloscopie virtuelle, les auteurs de [48] ont effectué une coloscopie virtuelle et une coloscopie comme test de référence sur 70 patients. Les résultats ont montré une sensibilité et une spécificité :

- de 75% et 90% pour les adénomes de taille supérieure à 1cm,
- de 66% et 63% pour les adénomes de taille comprise entre 5 et 9mm
- de 45% et 80% pour les adénomes de taille inférieure à 5mm.

Il est clair que la coloscopie virtuelle est une bonne méthode de diagnostic à partir du moment où les adénomes ont une taille supérieure à 1cm. De plus, cette méthode n'a pas besoin d'anesthésie et a un bon niveau d'acceptation par le patient, ce qui sont des avantages par rapport à la coloscopie. Il reste néanmoins ses faibles performances pour détecter les petits polypes et cela expose le patient à des radiations. Enfin, la coloscopie virtuelle ne permet aucunement l'ablation des polypes.

Au total, la coloscopie virtuelle est actuellement validée par la Haute Autorité de Santé comme une alternative diagnostique chez les patients en échec ou contre-indication ou refus de coloscopie.

1.2.4 Capsule Endoscopique ([CE](#))

La première [CE](#), la M2A, a été inventé par Paul Swain, Gavriel Iddan, Gavriel Meron and Arkady Glukhovsky [87] dans les années 1990, elle a été approuvée par le Food and Drug Administration (FDA) américaine pour être utilisé dans la visualisation

de l'intestin grêle en août 2001. Elle a ouvert de nouvelles opportunités notamment dans le diagnostic des pathologies de l'intestin grêle.

Il s'agit d'une pilule intégrant une caméra et un système de radio communication que le patient avale. Une fois avalée, la capsule évolue dans le tube digestif pendant environ 8 à 12 heures. Durant ce transit elle transmet les images du tube digestif vers une boîte externe (voir la figure 1.5) via une communication radio-fréquence. Une fois toutes les images acquises, un spécialiste les télécharge pour réaliser une analyse visuelle et déterminer si il y a des anomalies.



FIGURE 1.5 : Capsule endoscopique : PillCam

Comme tout dispositif médical ou procédure, la [CE](#), pour le côlon, présente des avantages et des inconvénients par rapport à la coloscopie traditionnelle.

Parmi les avantages citons :

- Dispositif très peu invasif.
- Période de jeûne souvent la même que pour l'endoscopie traditionnelle.
- Pas de période de récupération après la procédure.
- Réduction du risque de saignement causé par la procédure.
- Risque d'infection réduit.
- Pas besoin de sédation ou d'anesthésie.

Parmi les inconvénients citons :






- Obtention des résultats après 24 à 48 heures.
- Risque minime (<1%) d'impact de l'appareil dans tube digestif.
- Très exceptionnellement, une fausse route lors de l'ingestion peut amener à l'obstruction respiratoire.

- Les biopsies et l'ablation de polypes ne peuvent pas être effectuées pendant la procédure.
- Pas de nettoyage possible pendant la procédure. Le nettoyage est nécessaire dans le cas d'un résidu qui recouvre une anomalie intestinale.

1.3 CE disponibles dans le marché :

Aujourd'hui, différents types de CE sont disponibles sur le marché. Certaines capsules ne comportent qu'une caméra, d'autres deux, ou quatre. Plus de 75 millions de patients utilisent cette technologie chaque année. Les caractéristiques de ces capsules sont affichées dans le tableau 1.3.

TABLE 1.3 : Capsules endoscopiques disponibles sur le marché

Capsule	 PillCamColon 2 Medtronics	 EndoCapsule Olympus	 MiroCam Intromedic	 OMOM Jinshan	 CapsoCam CapsoVision
Largeur[mm]	31,5	26	24.5	27.9	-
Diamètre[mm]	11,6	11	10.8	13	-
Autonomie[h]	>10	>12	>11	>10	>15
Résolution	2x340x340	512x512	320x320	640x480	4x288x212
Fps	4-35[fps]	2[fps]	3[fps]	2[fps]	20[fps]
Angle de vue	172°	145°	170°	140°	360°

1.3.1 Medtronic : PillCam

La PillCamTM est le successeur de la première CE, la M2A. Medtronic a réalisé différentes capsules PillCam chacune conçue pour cibler un problème ou une section spécifique du tractus gastro-intestinal : maladie de Crohn, l'œsophage (rarement pratiquée), l'intestin grêle (principale indication) et le côlon (indication émergente). Actuellement, Medtronic est le leader du marché, avec plus de deux millions des patients qui ont utilisé une de leur CE pour l'analyse de l'intestin grêle, de l'œsophage et du côlon, soit plus de 125 000 patients par an. Une des capsules PillCam, la PillCam côlon 2 [7] est spécialisée pour l'analyse du côlon, voir ses caractéristiques dans la table 1.3. À l'aide de cette capsule les médecins peuvent obtenir une sensibilité et une spécificité acceptable de 68,8% et 81,3% pour identifier au moins un polype de 5mm ou plus [39]. Elle possède une autonomie énergétique maximale de dix heures,

ceci est juste pour inspecter le côlon, dans la plupart des cas la transmission est arrêtée avant d'arriver à la fin du côlon [16]. Les images acquises ont une résolution spatiale de 256x256 pixels, ce qui est faible par rapport à la coloscopie qui utilise des images d'au moins 1920x1080 pixels. Cette CE a une résolution temporelle de 4 à 25 images par seconde, ceci lui permet de transmettre et enregistrer entre 144 000 et 1 260 000 images dont la plupart ne reflètent aucune anomalie. Ce nombre important d'images rend l'analyse assez chronophage pour le gastro-entérologue.

1.3.2 Jinshan : OMOM

La capsule OMOM 2 [5] se caractérise par une résolution d'image de 640x480 pixels et une autonomie de plus de 10 heures. Cette capsule offre une autonomie suffisante pour l'inspection de l'intestin grêle [5]. Il n'existe pas de travaux scientifiques l'utilisant pour l'analyse du colon.

1.3.3 Olympus : EndoCapsule 10

Olympus propose l'EndoCapsule 10 [2], une CE pour l'analyse de l'intestin grêle, voir ses caractéristiques en table 1.3. Elle possède une autonomie énergétique minimale de 12 heures et une résolution spatiale de 512x512, grâce au traitement des images intégré dans la capsule qui élimine les images redondantes et non pertinentes, une réduction de 64% des images transmises est obtenue diminuant d'autant le temps nécessaire à leur analyse. Cette capsule a été conçue pour l'inspection de l'intestin grêle et son constructeur indique uniquement une utilisation possible pour cette partie de l'intestin [2].

1.3.4 CapsoVision : CapsoCam Plus

La capsule CapsoCam Plus [1], développée dans la Silicon Valley, se caractérise par une vue latérale panoramique à 360° de la muqueuse de l'intestin grêle, grâce aux quatre caméras latérales. Cette capsule offre une image panoramique de résolution de 1152x212 (quatre images de 288x212). Elle possède une autonomie énergétique minimale de 15 heures en stockant les images prises dans la capsule. Ces images sont téléchargées à la fin de la procédure à l'aide d'un module appelé CapsoAccess. Cette capsule offre une autonomie suffisante pour l'inspection de l'intestin grêle dans la plupart des cas, mais cette capsule ne permet pas l'analyse du côlon. Bien que

cette capsule offre une image panoramique à partir de l'union de quatre images, la résolution de chaque image est encore faible.

1.4 Limites actuelles de la CE :

TABLE 1.4 : Performances des CE

Métrique	Polypes	Polypes adénomes	CE
Sensibilité [%]	64	73	74
Spécificité [%]	84	79	-

Des études récentes [8], [44] indiquent qu'il est difficile de déterminer la véritable sensibilité de la CE en raison du manque d'une norme standard adéquate. Toutefois, les auteurs de [91] ont testé la CE et la coloscopie sur 328 patients ayant un âge moyen de 58,6 ans pour déterminer la sensibilité et spécificité de la CE par rapport à la coloscopie. Les résultats sont visibles dans le tableau 1.4. Dans leurs résultats, la CE a montré un bon degré d'acceptation par les patients, mais elle a montré un taux de détection inférieur à la coloscopie. Enfin, la capsule colique, comme la coloscopie virtuelle, ne permet aucunement la résection des adénomes. Autrement dit, un patient chez lequel un adénome a été identifié par capsule colique doit avoir une coloscopie pour sa résection. En pratique, la Haute Autorité de Santé a validé en France actuellement comme seule indication de la capsule colique l'échec de la coloscopie lorsque cet échec n'est pas lié à des difficultés de purge du côlon.

Dans [74], les auteurs ont réalisé une étude des progrès et défis techniques de la CE. Dans leurs résultats ils ont remarqué que par rapport aux endoscopes, la faible résolution spatiale et temporelle de l'image et l'autonomie limitent une utilisation plus large de la CE.

Les auteurs de [85] ont réalisé une étude des indicateurs de qualité intraprocédural d'endoscopie par capsule. Ces indicateurs sont le produit des limites techniques liées à la capsule et peuvent réduire la qualité des images. Leurs résultats ont permis d'identifier les limites techniques liées au système de la CE comme une résolution spatiale et temporelle faible et une autonomie insuffisante.

Des rapports publiés sur le taux d'erreur technique dans la pratique de la coloscopie [36], ont montré que la haute-définition des images permet d'obtenir un meilleur taux de détection des adénomes. La qualité des images joue un rôle crucial dans le diagnostic.

Lee et al. [63], dans leur travaux, ont mesuré le temps de trajet d'une capsule dans le tractus gastro-intestinal. Dans leur analyse ils ont utilisé une Wireless Motility Capsule (WMC) ingérable. La plage normale de temps de transit comprend les

éléments suivants : transit gastrique (2 à 5 heures), transit dans l'intestin grêle (2 à 6 heures), transit dans le côlon (10 à 59 heures) et transit de tout le tract gastro-intestinal (14 à 70 heures). L'autonomie nécessaire pour que la capsule assure l'intégralité du transit dans le côlon doit être de plus de 10 heures.

Dans [74], les auteurs ont identifié le goulot d'étranglement pour avoir une meilleure qualité d'images et autonomie dans cette nouvelle technologie qui est l'alimentation de la CE.

Sur la base de ces résultats, il convient de prendre des mesures pour augmenter la résolution spatiale et temporelle afin d'obtenir des images de meilleure qualité et une autonomie énergétique supérieure à 14 heures pour permettre le transit de tout le tract gastro-intestinal ou supérieur à 10 heures pour le transit dans le côlon.

Les constructeurs de capsule ont tenté de repousser ces limites :

- la PillCam Colon 2 a deux résolutions temporelles 4 et 35 images par seconde. Le changement de résolution temporelle s'opère automatiquement selon la région du côlon dans laquelle se situe la capsule, ce changement est un algorithme propriétaire, qui n'est pas publique, qui se base sur la connaissance des différentes régions anatomiques du colon. . Cette capsule a deux piles Energizer ou Renata 339, chaque pile a un diamètre de 9,5mm et une capacité de 51mAh pour 1,2 Volts. La valeur moyenne de la puissance s'élève à 7,3 mW et 255,5mW respectivement et la valeur moyenne du courant est de 3,8 mA et 133mA respectivement, ce qui offre une autonomie approximative de 14,47h et 0,41h respectivement. D'après Medtronic, les PillCam Colon 2 fonctionnent pendant au moins 10h.
- EndoCapsule intègre un traitement des images dans la capsule qui élimine les images redondantes et non pertinentes pour réduire de 64% les images transmises et avoir une autonomie d'au moins de 12 heures, une résolution spatiale de 512x512 pixels et une résolution temporelle de 2 images par seconde.
- CapsoCam possède une autonomie énergétique minimale de 15 heures grâce au stockage interne des images. Ces images sont téléchargées à la fin de la procédure à l'aide d'un module appelée CapsoAccess. Elle incorpore quatre caméras qui ensemble offrent une résolution spatiale de 1152x212 pixels et une résolution temporelle de 20 images par seconde.

Une solution pour aborder la problématique de l'autonomie énergétique pourrait être d'utiliser une pile de plus grande capacité. Il existe aujourd'hui plusieurs autres piles sur le marché qui peuvent offrir une capacité plus grande, voir table 1.5, jusqu'à 188mA au prix d'une augmentation de 2 millimètres de la longueur de la capsule.

TABLE 1.5 : Piles bouton dans le marché

Fabricant Model	Energizer EPX76	Sony 357	Duracell VAL-AG13	Energizer 339
Diamètre[mm]	11,25	9	11,6	9,5
Hauteur[mm]	5	5,4	5,4	1,45
Capacité[mAh]	188	165	155	51
Voltage[V]	1,55	1,55	1,55	1,2

En conclusion, la limite la plus important des capsules actuelles est en terme d'usage. Une capsule aujourd'hui réalise une fonction de **data-logger** qui permet la collecte des images et pas leurs interprétations. Du point de vue du dispositif médical elle peuvent remplacer la coloscopie pour l'exploration du côlon, mais elles n'intègrent aucun capacité de diagnostic.

1.5 L'avenir de la CE

La CE a révolutionné le diagnostic des pathologies de l'intestin grêle.

Malgré les progrès techniques, les images de CE sont toujours examinées manuellement. Ce processus prend beaucoup de temps et peut être sujet aux erreurs de diagnostic en raison de la limitation naturelle des capacités humaines.

Des méthodes pour améliorer la qualité des images pourraient contribuer à réduire les erreurs d'interprétation humaine, par exemple des méthodes qui améliorent le contraste [73], le débruitage des images [46], l'augmentation artificielle de la résolution temporelle [47], la reconstruction d'une image 3D à partir des images 2D [79], etc...

Ils existent aussi différents travaux pour estimer la localisation de la capsule en utilisant un réseau des capteurs magnétiques externes [90] ou à l'aide d'un logiciel basé sur l'estimation de mouvement de la caméra de la capsule [86]. D'autres travaux étudient la stabilisation du mouvement de la capsule à l'aide de trois bras miniatures [27] ou à l'aide d'un récipient extensible attaché à l'arrière de la capsule [43]. D'autres concernent la réalisation de la gestion des données provenant de la capsule et leur stockage à l'aide d'un serveur de type nuagique [37].

Une réelle rupture technologique serait l'intégration dans la capsule d'une méthode de détection automatique de polypes. Cela aurait comme premier impact de rendre la capsule apte à faire un diagnostic précoce, qui sera ensuite validé par la supervision d'un médecin. La détection automatique de polypes pourrait, en conséquence, contribuer à réduire le temps de révision requis par les spécialistes et les erreurs d'interprétation humaine.

Cette capsule deviendrait un dispositif de diagnostic précoce du CCR et pourrait être d'une grande utilité pour le suivi de patient à risque moyen et élevé, réduisant ainsi les risques liés à l'anesthésie de la coloscopie et à l'examen lui-même.

Des méthodes de détection automatique des polypes des ulcères ou d'hémorragies, ont fait leur apparition au début des années 2000 mais aucune des ces méthodes n'a été conçus pour une intégration à l'intérieur d'une capsule endoscopique.

Synthèse et problématique

Dans ce chapitre nous avons introduit la prévention du [CCR](#) comme un enjeu majeur de santé publique. Nous avons montré l'importance et le rôle des politiques de dépistage, notamment avec une nécessité d'augmenter l'acceptation des tests par les patients. La détection des adénomes à un stade précoce avant qu'ils ne deviennent cancéreux permet de réduire significativement la survenue du [CCR](#).

Nous avons décrit le système de dépistage en France qui commence par un test, le [FIT](#) suivi, si le test est positif, par une coloscopie. En cas d'échec de la coloscopie, ils existent d'autres méthodes comme la Coloscopie Virtuelle et la Capsule Endoscopique [CE](#).

Nous avons décrit le test [FIT](#), qui est insuffisamment utilisé malgré sa facilité d'usage et sa haute spécificité.

La coloscopie a été présentée comme le test de référence utilisé par les gastro-entérologues. Bien que son taux de détection soit élevé et qu'elle permette de façon unique la résection des polypes, la coloscopie reste une méthode invasive et inconfortable. Elle ne peut pas être utilisée comme un dispositif de dépistage à l'échelle individuelle car son acceptabilité est limitée par l'obligation d'une phase de nettoyage qui oblige la personne à se purger.

Nous avons décrit la coloscopie virtuelle comme une méthode moins invasive que la coloscopie qui ne requiert pas d'anesthésie, ce qui augmente le taux d'adhésion des patients pour le diagnostic. Cette méthode n'est pas performante pour détecter des polypes plus petits qu'un centimètre, et n'est donc pas indiquée pour le dépistage, elle expose le patient à des radiations et ne permet pas l'ablation des polypes.

Enfin, nous avons introduit la [CE](#) comme une méthode avec un bon degré d'acceptation des patients, mais avec un taux de détection inférieur à la coloscopie. La [CE](#) présente quelques limites et faiblesses liées à la résolution spatiale et temporelle des images, à son autonomie énergétique et au nombre d'images transmises à analyser par le gastro-entérologue.

A partir de ces éléments, mes travaux de thèse s'orientent vers la conception d'un nouveau dispositif médical pour le diagnostic précoce du [CCR](#) avec un degré d'acceptation qui sera plus important que celui d'une coloscopie, il n'y aura pas d'hospitalisation ni d'anesthésie. Le diagnostic précoce intégré dans une [CE](#) avec

une capacité de détection automatique des polypes nous permettra de dépasser le limite du test FIT et de réduire les effets secondaires de la coloscopie et de la coloscopie virtuelle.

Ce nouveau dispositif devra améliorer les conditions de confort du patient et être sans risque pour la santé du patient. Il devra détecter des polypes plus petits qu'un centimètre. En prenant en compte toutes les limites et faiblesses des CE existantes, liées à la faible résolution spatiale des images, à l'autonomie énergétique et au nombre trop important d'images transmises, ce nouveau dispositif médical pour le diagnostic précoce du CCR pourra réduire sa mortalité, son incidence et sa mortalité.

Nous proposons un nouveau paradigme de capsule endoscopique : la **Capsule Endoscopique Intelligente (iCE)**. Elle sera capable de détecter des polypes in situ permettant de ne transmettre que les images évoquant un polype. Grâce à la diminution des images transmises l'autonomie de la capsule augmentera à plus de 12 heures tout en assurant la même résolution spatiale et temporelle des coloscopes : 25 images de 1920x1080 pixels par seconde.

Comment rendre la CE intelligente ? En intégrant un **Circuit de Vision Intelligent (Smart Vision System on Chip (SVSoC))** qui contient une chaîne de traitement pour la détection des polypes dans un système sur puce.

Nous proposons une capsule qui offre une qualité des images au niveau du coloscope. Ici, les contraintes sont liées à :

- l'application en temps réel à 25 images par seconde ;
- une surface limitée à celle de la capsule (surface de $8 \times 8 \text{ mm}^2$ pour une capsule de diamètre de 11,6 mm) ;
- un diamètre autour de 11,6 mm et une largeur autour de 31,5 mm ;
- un budget énergétique limité à deux piles avec une capacité d'alimentation total de 376 mA ;
- une autonomie supérieure à 10 heures ;
- une seule usage.

Ce nouveau paradigme ouvre la voie à l'utilisation de la iCE pas seulement comme une alternatif à la coloscopie mais aussi comme une nouvelle étape intermédiaire entre le dépistage réalisé par le FIT et le diagnostic fait par le médecin avec la coloscopie. Les travaux présentés ici se sont focalisés sur le CCR, mais il est tout à fait envisageable de considérer d'autres pathologies et de réaliser une iCE pour chaque pathologie.

Deuxième partie

DÉTECTION AUTOMATIQUE
DES POLYPES

Introduction	33
2 Contexte et état de l'art	35
2.1 Méthodes pour la reconnaissance automatique des polypes	37
2.1.1 Méthodes de description et classification basées sur des algorithmes de traitement d'image	40
2.1.2 Méthodes utilisant des algorithmes de traitement d'image pour les caractéristiques et des algorithmes d'apprentissage automatique pour la classification	41
2.1.3 Méthodes de description et classification par apprentissage automatique :	43
2.2 Analyse	44
3 Classification des polypes par arbres flous	47
3.1 Extraction des ROI	49
3.2 Description des ROI	50
3.3 Apprentissage des arbres flous	52
3.3.1 Discrétisation des attributs	53
3.3.2 Mesure d'information H	54
3.3.3 Stratégie de Partition P	54
3.3.4 Critère d'arrêt T	55
3.4 Test de la classification des ROI avec les arbres flous	55
3.5 Classification des ROI avec une forêt d'arbres flous	58
3.6 Discussion	59
4 Extraction automatique des ROI	61
4.1 Experimentations	63
4.2 Extraction et suivi des ROI pour la détection des polypes	67
4.2.1 Suivi des ROI	69
4.2.2 Expérimentations	70
4.3 Discussion	72

5	Entrainement du classifieur en utilisant les ROI extraites par la chaîne de traitement	73
5.1	Expérimentations	74
5.2	Un approche de Boosting pour augmenter la performance de détection	77
5.2.1	Expérimentations	78
5.3	Mesures au niveau polype et patient	81
	Synthèse et discussion	83

Introduction

Le **SVSoC** qui doit être réalisé dans le cadre de ces travaux et qui est nécessaire à la détection précoce du **CCR** va intégrer en son sein un certain degré d'intelligence afin d'avoir **la capacité de reconnaître des polypes**. Les travaux que nous avons menés ont permis d'étudier comment réaliser ce **SVSoC** afin de faire cette reconnaissance en temps réel à partir des images acquises par la caméra intégrée dans une **CE**.

L'**imagerie numérique** a profondément changé la médecine et ses méthodes depuis plus de deux décennies, ceci est d'autant plus vrai pour la gastro-entérologie où l'image est la modalité principale pour réaliser un diagnostic. Parmi ces changements est apparue la capacité de détection automatique des polypes qui a été un domaine de recherche très actif durant ces 25 dernières années [25]. L'aide apportée par ces traitements automatiques permet d'améliorer les conditions de travail du médecin, par exemple en sélectionnant les images d'intérêts pour permettre la diminution du nombre d'images analysées, et par la même la diminution du temps d'analyse des images prises par l'endoscope ou la **CE**.

Dans cette partie, un premier chapitre propose un état de l'art sur les méthodes automatiques pour la reconnaissance des polypes à partir d'images de coloscopies et de **CE**. Suite à l'analyse de cet état de l'art, une première contribution scientifique, la conception d'un classifieur basé sur des arbres flous regroupés dans une forêt, est présentée. L'étude de l'intégration de ce classifieur dans une chaîne complète de traitement avec extraction automatique de régions d'intérêt (**ROI**) dans l'image, est l'objet du chapitre qui suit. De cette étude découle la seconde contribution de la thèse, une nouvelle approche de suivi pour améliorer l'extraction automatique des **ROI** et son intégration dans notre chaîne de traitement. Une synthèse et une discussion sur la méthode et les résultats obtenus clos cette partie.

En raison du contexte transdisciplinaire, la collaboration entre cliniciens et scientifiques est essentielle [89] pour la conception et l'intégration de cette capacité de détection au sein d'une **CE**. Cette collaboration a été mise en place dans le cadre de ces travaux entre le laboratoire LIP6 et le service d'endoscopie digestive de l'hôpital Saint Antoine. Elle nous a permis d'analyser et comprendre les enjeux et les difficultés à la fois dans le domaine médical et le domaine scientifique. Cette double analyse

nous a permis d'identifier quels sont les verrous à lever pour obtenir un dispositif biomédical efficace. Ces travaux ont été de plus soutenus par l'institut universitaire d'ingénierie pour la santé de Sorbonne Université, l'IUIS, dont l'une des missions est de favoriser les échanges entre chercheurs et cliniciens.

Chapitre 2

Contexte et état de l'art

*Sometimes science is more art than science,
Morty. Lot of people don't get that.*

Rick Sanchez

Un gastro-entérologue doit analyser en temps réel dans le cas d'une coloscopie, ou en différé si l'examen provient d'une [CE](#), les images du côlon. Ces images montrent les parois du côlon, mais aussi parfois des résidus de matière fécale ou de l'eau. Ces derniers occultent une partie de la paroi et rendent plus difficile l'énoncé de conclusions liées à l'examen. Pour venir en aide au médecin lors de l'examen de ces images plusieurs méthodes d'analyse automatiques ont été développées. Elles permettent soit d'assister le médecin pendant une coloscopie, pour augmenter son taux de reconnaissance, soit de l'assister lors de la lecture d'une vidéo provenant d'une [CE](#). Une synthèse de ces méthodes est accessible dans les articles [51] et [25].

Dans [51], les auteurs ont fait une analyse des contributions existantes pour la reconnaissance automatique des polypes en prenant en compte les aspects médicaux et scientifiques. Ils ont mis en évidence le manque d'une [norme de référence \(gold standard\)](#) pour évaluer les performances des différentes méthodes. Cette norme de référence pour l'image est généralement constituée d'une base de données annotées prenant en compte une grande variété de scénarios possibles. Des métriques pertinentes permettant les comparaisons des différentes méthodes sont aussi généralement définies. Pour répondre à ce besoin Bernal & al.[25] ont proposée quatre bases de données annotées lors des challenges [Gastrointestinal Image Analysis \(GIANA\)](#), pour l'évaluation des performances de reconnaissance automatique des polypes. Ils ont organisé ce challenge dans la conférence internationale [Medical Image Computing and Computer Assisted Intervention \(MICCAI\)](#) entre 2015 et 2018. Ces bases de données ont été créées et mises à la disposition de la communauté scientifique. Elles ont pour nom **ASU-Mayo Clinic Colonoscopy Database**, **CVC-CLINIC**, **ETIS-LARIB** et **CVC-VideoClinicDB**.

TABLE 2.1 : Détails de ASU-Mayo Clinic Colonoscopy Database [89].

Base d'apprentissage					
Film	Images (avec polype)	Résolution	Film	Images (avec polype)	Résolution
1	682(0)	712x480	11	324(245)	1920x1080
2	838(0)	712x480	12	910(910)	1920x1080
3	769(0)	712x480	13	519(374)	1920x1080
4	712(0)	712x480	14	501(391)	856x480
5	1843(0)	712x480	15	1200(1106)	856x480
6	1925(0)	712x480	16	339(209)	1920x1080
7	1550(0)	712x480	17	418(234)	856x480
8	1740(0)	712x480	18	259(189)	1920x1080
9	1802(0)	712x480	19	616(235)	1920x1080
10	1639(0)	712x480	20	410(385)	856x480

- **ASU-Mayo Clinic Colonoscopy Database [89]** : Cette base de données est constituée de 38 vidéos, 20 pour l'apprentissage et 18 pour le test, de coloscopies annotées réalisées par le département de gastro-entérologie de la clinique Mayo en Arizona aux Etats Unis. Dans cette base plusieurs résolutions spatiale d'images sont fournies : 712 x 480, 856x480 et 1920x1080. Seule la base d'apprentissage de cette base de donnée est publique, ses détails sont listés dans le tableau 2.1.
- **CVC-CLINIC [23]** : Cette base contient 31 vidéos pour un total de 612 images avec d'une résolution 388x284. Il y a 31 polypes différents soit un polype par vidéo.
- **ETIS-LARIB** : La base est constituée de 34 vidéos pour un total de 196 images avec une résolution de 1225x996. Dans cette base il existe 44 polypes différents.
- **CVC-VideoClinicDB** : cette base est constituée de 18 vidéos pour l'apprentissage et de 18 vidéos pour le test. Chaque vidéo contient des images d'une résolution spatiale de 384x288 pixels. Ces vidéos ont été extraites à partir d'explorations coloscopiques réalisées à la Clinique Hospitalière de Barcelone, en Espagne. Cette base de données vise à couvrir différents types des polypes avec pour chaque polype des points de vues différents. L'annotation de cette base est faite en suivant la classification de Paris [22], visible dans le tableau 2.2, qui permet de différencier les polypes en fonction de leur forme. Seule la base d'apprentissage de cette base de donnée est publique, ses détails sont listés dans le tableau 2.3.

Pour notre analyse et comparaison avec l'état de l'art nous avons utilisé seulement les bases de données d'apprentissage **ASU-Mayo Clinic Colonoscopy** et

TABLE 2.2 : Classification de Paris

Typ 0-I - Forme polypoïde	
0-Ip	lésion en relief, pédiculé
0-Is	lésion en relief, sessile
Typ 0-II - Forme plane	
0-IIa	lésion plane légèrement surélevée
0-IIb	lésion totalement plane
0-IIc	lésion plane légèrement déprimée
Typ 0-III - Forme ulcérée	
0-III	lésion creusante ou ulcérée

CVC-VideoClinicDB. Les bases de test ne sont pas publiques.

TABLE 2.3 : Détails de la base d'apprentissage CVC-VideoClinicDB [25, 20].

Base d'apprentissage					
Film	Images (avec polype)	Paris	Film	Images (avec polype)	Paris
1	498(386)	0-Is	10	840(762)	0-IIa
2	773(597)	0-Is	11	500(370)	0-Is
3	972(819)	0-Is	12	385(261)	0-IIa
4	390(350)	0-Is	13	624(620)	0-Is
5	490(412)	0-Is	14	2060(2015)	0-Is
6	857(522)	0-Ip	15	575(360)	0-Is
7	441(338)	0-Is	16	371(366)	0-Is
8	449(405)	0-IIa	17	797(651)	0-IIa
9	551(532)	0-Ip	18	381(259)	0-Ip

2.1 Méthodes pour la reconnaissance automatique des polypes

La **reconnaissance automatique des polypes** est constituée d'une première étape d'extraction des caractéristiques de l'image, sur la globalité de l'image ou sur des régions de l'image. Cette étape est suivie d'une seconde étape de classification de l'image à partir de ces caractéristiques. Pour représenter les caractéristiques des descripteurs sont calculés, ces descripteurs dépendent des caractéristiques recherchées : forme, couleur, texture, etc,

Pour réaliser ces étapes il est possible de faire appel à des méthodes de traitement d'image et/ou à des algorithmes de classification automatique.

Le traitement d'image est un domaine qui a vu son émergence dans les années

1950 et dont l'objectif est de manipuler à différents niveaux d'abstraction les objets présents dans une image. Dans le cadre de ces travaux nous avons considéré un traitement bas-niveau c'est à dire au niveau du pixel, les traitements permettent à ce niveau : la modélisation, la reconstruction, la restauration ou l'analyse d'image [15].

L'apprentissage automatique est un domaine de l'intelligence artificielle qui est apparu dans les années 1950, le terme vient de l'anglais "Machine Learning" qui a été introduit la première fois en 1959 dans un article de Arthur Samuel [81] un des pionniers de l'intelligence artificielle. L'apprentissage automatique est basé sur l'analyse statistique des données afin de résoudre des tâches comme la classification des données. Il existe deux phases dans l'apprentissage automatique, une première phase d'apprentissage à proprement parler où à l'algorithme va apprendre à partir des données d'entrée à réaliser la tâche de reconnaissance et une seconde phase de reconnaissance dans laquelle l'algorithme, ayant appris à réaliser la tâche de reconnaissance, va jouer pleinement son rôle en classant les images d'entrée. Il existe plusieurs techniques d'apprentissage automatique : les arbres de décision binaires, les arbres flous, les machines à vecteurs support, les réseaux de neurones artificiels, etc

Un arbre de décision est une hiérarchie de tests d'éléments appelés attributs, ce concept a été introduit par J. Morgan en 1963 [70]. Les nœuds de l'arbre représentent les attributs et les arcs représentent les tests comme nous pouvons le voir dans la figure 2.1. La hiérarchie commence par un nœud principal contenant l'attribut le plus discriminant, puis l'arbre est parcouru à l'aide des tests présents sur les arcs. Le parcours de l'arbre traverse plusieurs nœuds représentant des attributs plus ou moins discriminants, jusqu'à arriver à une feuille qui représente la décision.

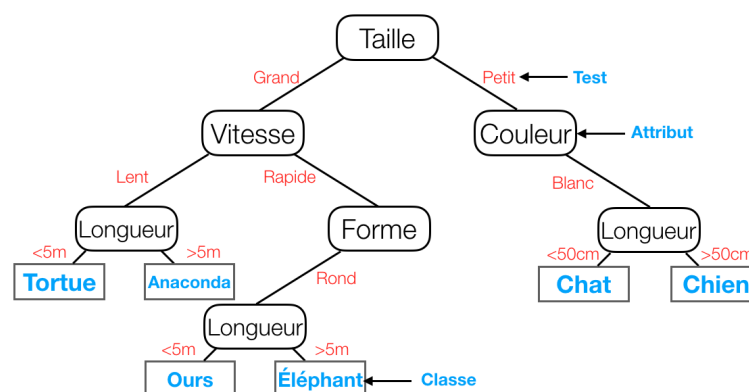


FIGURE 2.1 : Arbre de décision

Un arbre de décision flou est une hiérarchie de calcul d'appartenance, c'est la fusion du concept d'arbre de décision et de la logique floue décrite la première fois en 1965 par L. Zadeh [97]. Les nœuds représentent les attributs et les arcs représentent des fonctions d'appartenance. Ces fonctions d'appartenance cascadiées indiquent au

niveau des feuilles une probabilité de classe. Contrairement aux arbres de décision binaire, ici toutes les branches de l'arbre sont parcourues.

Dans les machines à vecteur support ou séparateurs à vaste marge (en anglais Support Vector Machine), introduites en 1992 pour l'apprentissage automatique [26], il est recherché l'hyperplan de marge optimale permettant de séparer correctement un ensemble des données. L'hyperplan est défini au moyen d'une fonction noyau qui permet de projeter les données dans un espace avec plus de dimensions comme cela est représenté dans la figure 2.2.

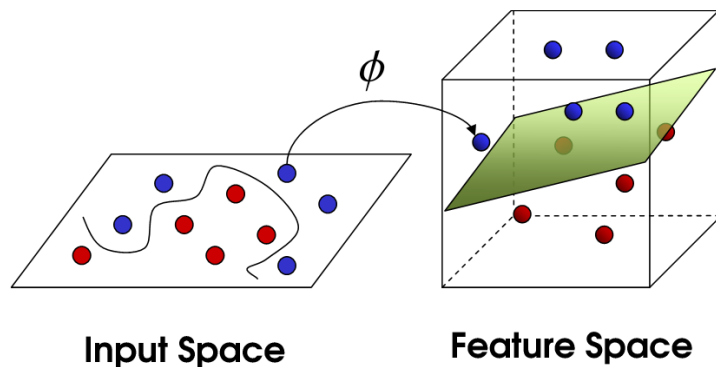


FIGURE 2.2 : Machines à vecteurs de support

Les réseaux de neurones artificiels sont des systèmes inspirés des réseaux de neurones biologiques qui constituent le cerveau, c'est en 1943 que McCulloch et Pitts les introduisent [60]. Dans ce système, un ensemble d'unités appelés neurones sont interconnectées à l'aide de liens appelés synapses. Un neurone va participer à exciter ou à inhiber les neurones auxquels il est connecté. Plusieurs modèles d'organisation existent, le plus courant est celui qui structure le réseau de neurones en couches de neurones, l'évaluation de l'excitation ou de l'inhibition des neurones est réalisée couche par couche, les réseaux de neurones appelés Perceptrons Multi-Couches [82] mettent en œuvre cette organisation.

Un autre modèle de réseaux de neurones a pris une place importante depuis 2012 lors du challenge ILSVRC ¹. Il s'agit des réseaux de neurones profonds possédant plusieurs millions de synapses [56]. Le nom couramment utilisé est Deep Learning. Dans ces réseaux les neurones sont organisés en couches. Les couches peuvent être reliées entre elles par des connexions locales afin de réaliser des convolutions qui filtreront les données, une couche finale sera quand à elle totalement connectée. Dans ce cas nous parlerons de réseaux de neurones convolutionnels dont le premier modèle opérationnel fut introduit par Y. Lecun en 1989 [62]. Les réseaux de neurones

¹Imagenet Large Scale Visual Recognition Challenge

convolutionnels sont des réseaux dont les couches intermédiaires, dite couches de convolution, correspondent à filtres capables d'extraire des descripteurs complexes à partir de l'image.

Comme nous le constatons il existe plusieurs méthodes permettant de faire de la reconnaissance automatique. Parmi ces méthodes, celles pour la reconnaissance de polypes peuvent être regroupées en 3 groupes en fonction de la manière d'extraire les caractéristiques et de la manière de réaliser la classification :

1. il y a les méthodes réalisant l'extraction des caractéristiques et la classification à partir d'algorithmes de traitement d'image bas ou moyen niveau ;
2. il y a les méthodes réalisant l'extraction en utilisant des algorithmes de traitement d'image bas ou moyen niveau et la classification en utilisant des algorithmes d'apprentissage automatique ;
3. il y a les méthodes réalisant l'extraction des caractéristiques et la classification en utilisant des algorithmes d'apprentissage automatique.

L'état de l'art de ces trois méthodes fait l'objet de la prochaine partie.

2.1.1 Méthodes de description et classification basées sur des algorithmes de traitement d'image

Les traitements dans ce groupe exploitent des algorithmes simples de traitement d'images pour extraire des régions suspectes et/ou pour extraire des descripteurs. Parmi les travaux existants nous pouvons citer :

- Les travaux de **Karargyris et al., 2011[54]** pour la détection automatique des ulcères et des polypes afin de réduire le temps d'analyse par le médecin des vidéos produites par une [CE](#). Dans la chaîne de traitement, les auteurs appliquent d'abord un filtre Log-Gabor et l'algorithme Susan pour extraire les contours dans l'image. Ensuite les centres de courbure entre chaque couple de points de contour sont accumulés. Puis l'image est segmentée à partir des points où il y a le plus d'accumulations, comme nous pouvons le voir dans l'image 2.3. Des descripteurs sont extraits de chaque segment : la courbure et l'excentricité. Ces descripteurs sont classés simplement en utilisant deux seuils. Les auteurs ont mentionné que la complexité de calcul du détecteur de contours Susan et le calcul des centres de courbure rendent impossible l'utilisation de cette chaîne de traitement en temps réel, mais une mise en œuvre matérielle pourrait éventuellement rendre son utilisation en temps réel possible.
- Dans la méthode décrite par **Bernal et al., 2012[24]** l'image originale est segmentée avec une technique utilisant le gradient de l'intensité de l'image. Il

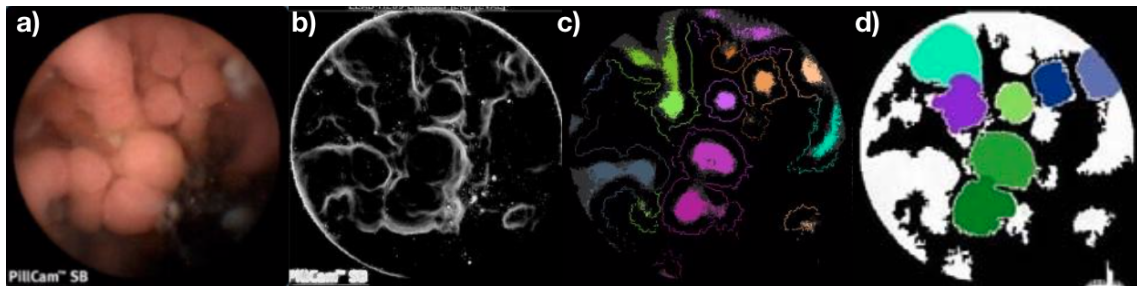


FIGURE 2.3 : a) image originale, b) image de contours, c) accumulation des centres de courbure et d) segmentation de l'image.

a été constaté que l'intensité est plus importante dans la vallée d'un polype. L'accumulation des gradients de l'intensité permet donc d'indiquer la position du polype, comme nous le voyons dans la figure 2.4. Cette technique s'appelle Sector Accumulation-Depth of Valleys Accumulation (SA-DOVA).

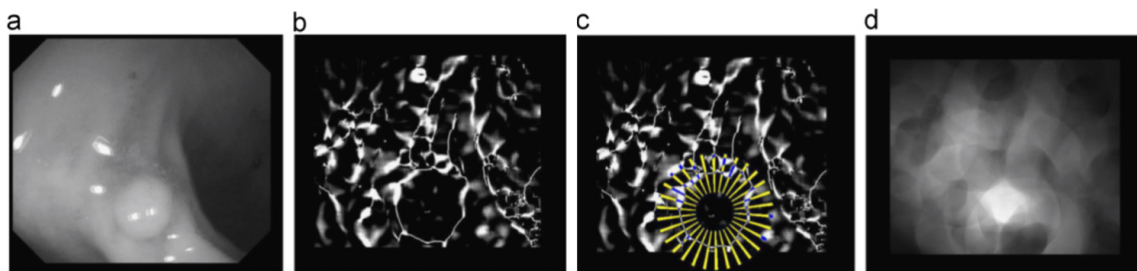


FIGURE 2.4 : a) image originale, b) image de contours, c) gradients d'intensité d) accumulation des gradients d'intensité.

2.1.2 Méthodes utilisant des algorithmes de traitement d'image pour les caractéristiques et des algorithmes d'apprentissage automatique pour la classification

Dans ce groupe divers travaux existent, parmi ceux-ci nous avons sélectionné les travaux suivants :

- Les travaux de **Mamonov & Figueiredo, 2014 [67]** qui permettent d'extraire des descripteurs grâce à une analyse de la géométrie et de la texture de l'image en utilisant l'algorithme de Buades [30], cet algorithme calcule l'inertie et l'excentricité comme descripteurs. La classification des descripteurs est faite avec un arbre de décision binaire.
- Les travaux de **Kodogiannis et al., 2007 [55]** qui définissent une méthode nommée CoLD (Colorectal cancer Lesions Detector). Ici l'image est représentée

en deux espaces de couleurs, [Red Green Blue \(RGB\)](#) et [Hue Saturation Value \(HSV\)](#) soit 6 couleurs au total. L'extraction des descripteurs de texture est faite sur les 6 canaux de couleur avec l'algorithme [Local Binary Pattern \(LBP\)](#). L'algorithme [LBP](#) compare le niveau de luminance d'un pixel avec les niveaux de ses voisins, chaque comparaison donne un bit "0" ou "1", l'ensemble des bits donnent un motif, ce qui donne information sur la texture d'un pixel dans un image, voir figure 2.5. Comme chaque pixel a 8 voisins il y aura 256 motifs différents de texture. L'information de texture de tous les pixels est stockée dans un histogramme à partir duquel il est possible de calculer des descripteurs statistiques de la texture de l'image. Le classifieur utilisé est un [Artificial Neural Network \(ANN\)](#) basé sur la logique floue appelée [Adaptive Fuzzy Logic System \(AFLS\)](#).

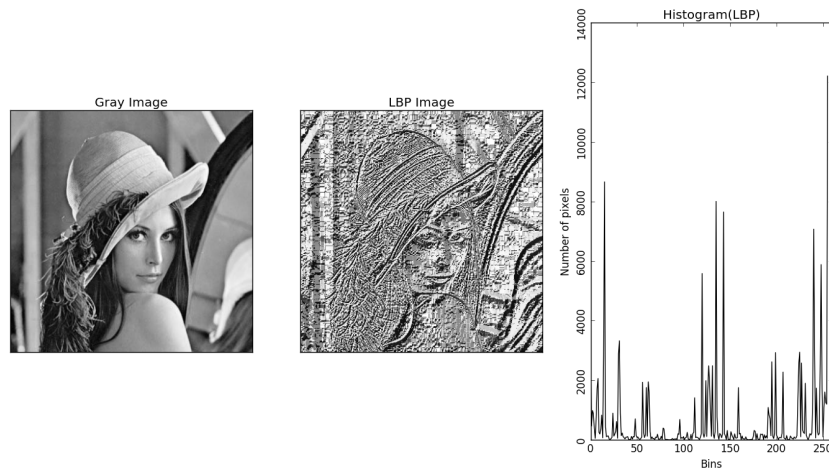


FIGURE 2.5 : Algorithme de Local Binary Pattern

- Les travaux **Li & Meng, 2011[66]** proposent une méthode qui consiste à extraire des descripteurs de texture à partir d'une matrice de co-occurrence calculée sur différentes sous-images extraites de l'image originale par une transformée en ondelettes. La matrice de co-occurrence est calculée en utilisant une variation de l'algorithme [LBP](#) nommé Uniform Binary Pattern. Cet algorithme utilise 10 motifs au lieu de 256 pour coder l'information de texture à partir de la comparaison du niveau de luminance d'un pixel avec les niveaux de ses voisins avec l'objectif de réduire la taille de l'histogramme, voir figure 2.6. Le classifieur utilisé est une Support Vector Machine avec un noyau RBF (Radial Basis Function).
- Dans les travaux de **David et al., 2013 [34]** les auteurs calculent tout d'abord l'histogramme des gradients orientés [Histogram of Oriented Gradients \(HOG\)](#)

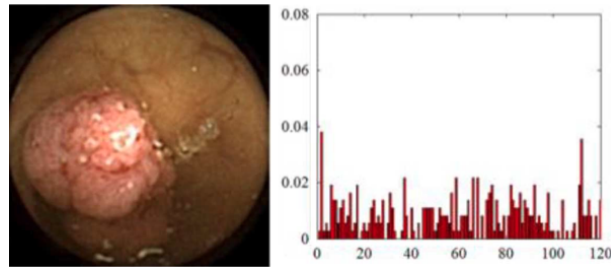


FIGURE 2.6 : Algorithme Uniform Binary Pattern

à partir de l'image originale. Ensuite, des descripteurs de forme sont extraits sur l'histogramme pour entraîner un ANN de type Perceptrons Multi Couches.

- les auteurs dans **Romain et al., 2013 [78]** proposent une méthode de détection automatique des polypes où des régions sont extraites avec la transformée de Hough. Chaque région est caractérisé par des descripteurs de texture extraits avec l'aide d'une méthode de calcul de matrice de co-occurrence. Finalement les descripteurs de chaque région sont classés avec une forêt d'arbres binaires de type Adaboost.

2.1.3 Méthodes de description et classification par apprentissage automatique :

Ce sont les méthodes dont l'extraction des descripteurs et la classification sont réalisés par apprentissage automatique. Comme dans plusieurs domaines aujourd'hui, ici les **Convolutional Neural Network (CNN)** sont utilisés pour détecter les polypes.

Parmi les travaux existants nous pouvons citer :

- Dans l'article [72, 25] les auteurs proposent un CNN composé de trois parties pour la détection automatique des polypes sur des images de CE. Une première partie réalise un sous-échantillonnage à l'aide de couches de convolution et de couches appelées Max-Pooling, une seconde partie réalise un sur-échantillonnage avec des couches de convolution et des couches appelées Up-sampling et une dernière partie totalement connectée classe les descripteurs extraits à l'aide des deux premières parties. Les auteurs ont participé au challenge de détection de polypes de la conférence MICCAI en 2015.
- Dans l'article [50, 25] les auteurs utilisent AlexNet [56], un réseau de neurone convolutionnel pré-entraîné à l'aide la base des images ILSVRC 2012 [80]. La taille de la couche d'entrée a été modifiée pour s'adapter à des images de 96x96 pixels, et les auteurs ont réduit la taille des noyaux des deux premières couches de Max-Pooling de 3 à 2. A la sortie du réseau, la dernière couche a été modifiée afin d'avoir seulement deux sorties : polype, non polype. Les auteurs

ont augmenté la base de données en utilisant des techniques de rotation, de symétrie, de redimensionnement et de recadrage sur des images de la base de données originale. Ils ont appelé leur réseau de neurones convolutionnel CaffeNet. Ils ont participé au challenge de détection des polypes de la conférence MICCAI 2015.

- Dans l'article [99, 25], les auteurs proposent trois réseaux de neurones convolutionnels qui ont été entraînés pour trois différentes échelles d'image. La dernière couche de chaque réseau de neurones convolutionnel a été éliminée pour permettre leur agrégation. Cette dernière est réalisée à l'aide d'un perceptrons multi-couches entraîné séparément. Ils ont participé au challenge de détection de polypes de la conférence MICCAI 2015.
- Dans l'article [94], les auteurs proposent un réseau de neurones convolutionnel 3D pour la détection automatique des polypes. Ce réseau analyse l'information spatio-temporelle d'un segment de vidéo d'une coloscopie avec une carte des probabilités. Il est capable d'extraire des descripteurs spatio-temporels plus représentatifs que des descripteurs spatiaux. La méthode a été évaluée avec la base de données ASU-Mayo [25].
- Dans l'article [98], les auteurs présentent un nouveau paradigme de réseau de neurones convolutionnel basé sur la régression temporelle pour la détection automatique des polypes en temps réel pendant la coloscopie. Leur méthode est composée de deux parties : une première partie est constituée d'un réseau pré-entraîné, ResYOLO, et une seconde partie est composée d'une régression temporelle réalisée avec des opérateurs de convolution nommés [Efficient Convolution Operators \(ECO\)](#). La méthode a été évaluée avec la base de données cliniques ASU-Mayo [25].

2.2 Analyse

Le tableau 2.4 fait une synthèse des différents travaux présentés pour la reconnaissance automatique des polypes en indiquant la source des images, les méthodes utilisées et les performances obtenues.

Toutes les méthodes présentées sont des travaux de recherches dont les expériences ont été menées sur des bases de données d'images acquises préalablement, soit lors d'une coloscopie, soit lors de l'utilisation d'une [CE](#). Aucune de ces méthodes n'a été testée durant un examen ce qui est important si nous désirons les utiliser comme aide au diagnostic durant la coloscopie par exemple. Dans ce cas il est aussi nécessaire que la méthode choisie puisse être exécutée en temps réel, avec

TABLE 2.4 : État de l'art : méthodes de détection automatique des polypes

Authors, année	DB Images (polypes)	Source	Descripteurs (méthode)	Sensi- bilité	Spéci- ficité
Méthodes de description et de classification par traitement d'image					
[53]	50 (10)	CE	Forme, texture	100	67,5
[54]	100	Coloscopie	Forme, texture	96,2	70,2
[24]	300 (300)	Coloscopie	Forme	89,0	98,0
[67]	18900 (230)	CE	Forme (Buades)	47	90,0
Méthodes de classification par apprentissage automatique					
[55]	140 (70)	CE	Texture(LBP)	97	94
[66]	1200	CE	Texture(LBP)	91,6	NA
[78]	1500 (300)	Coloscopie	Texture(Ocurrence)	91,0	95,2
[34]	30540	Coloscopie	Forme(HOG)	80,0	65,0
Méthodes de description et de classification par apprentissage automatique					
[25]	ASU-Mayo	Coloscopie	CNN	71,4	94,4
[25]	ASU-Mayo	Coloscopie	CNN	51,5	98,3
[99]	CVC-Clinic	Coloscopie	CNN	52,8	NA
[94]	ASU-Mayo	Coloscopie	CNN	71,0	NA
[98]	ASU-Mayo	Coloscopie	CNN	71,6	97

au moins une latence de traitement par image de 40 ms. Cette latence correspond à une cadence vidéo standard de 25 images par seconde.

Toutes ces méthodes ont été conçues pour être exécutées sur un ordinateur externe sans prise en compte de contraintes de latence, de consommation ou de surface, telle que celles qui existent pour une intégration dans une [CE](#).

La plupart des méthodes ont été testées sur des petites bases de données, sauf pour la méthode de Mamonov et al. [67] qui par contre présente une faible sensibilité de détection et la méthode de David et al. [34] qui présente une faible spécificité.

Comme le montre les travaux de Li & Meng [66] et Kodogiannis et al. [55] les descripteurs de texture sont très utiles pour la caractérisation des polypes. Dans ces travaux, les auteurs ont utilisé un réseau de neurones qui montre un bon taux de détection. De manière générale, bien que les méthodes utilisant les réseaux de neurones convolutionnels arrivent à de très bonnes performances en terme de taux de détection, il utilise tous une quantité de neurones et de synapses, cette grande nombre rends difficile leur intégration dans un système sur puce avec une surface de $8 * 8mm^2$ qui a une budget énergétique limitée, nous parlons ici de plus de 4 millions de synapses pour le réseau GoogleNet [88] qui est parmi les plus petits. Les auteurs de [4] ont fait un circuit qui a une consommation de 879 mW (pour une tension de 1 Volt) pour une itération de GoogleNet avec un devis de 23,6 images par seconde. Cette consommation représente une autonomie estimée de 0,43 heures pour un budget énergétique de 376mAh, ce qui n'est pas envisageable pour réaliser le trajet de tout le tract intestinal.

Dans les travaux de Romain et al. [78], ils proposent une méthode pour la reconnaissance automatique des polypes envisagé pour être intégrée dans une CE. Dans cette méthode la classification est réalisée à partir des régions susceptibles de contenir un polype. Ces régions sont extraites à l'aide d'algorithmes de traitement d'images et sont caractérisées par des descripteurs de texture. Le classifieur a été entraîné avec des régions extraites à partir d'une base de données de 300 images avec un polype et 1200 images sans polype provenant de coloscopies. Cette méthode obtient de bonnes performances, une sensibilité de 91% et une spécificité de 95% pour la reconnaissance des polypes.

C'est donc tout naturellement que je suis parti de cette méthode dans le cadre de mes travaux de thèse. D'autant plus que ces travaux ont été réalisés en collaboration avec l'équipe Syel du LIP6 dans laquelle j'ai effectué ma thèse.

J'ai pu observer que la limite principale de cette méthode est due au fait que les bonnes performances de reconnaissance d'un polype dépendent fortement de la capacité à bien extraire la région susceptible de contenir un polype. Les travaux réalisés dans cette méthode se sont d'avantages focalisés sur le classifieur. L'avantage de cette approche est le bas niveau de complexité du calcul propre aux arbres binaires qui rend faisable la mise en œuvre de ce classifieur dans une CE.

Les performances de reconnaissance sont toutefois à relativiser, les mesures ayant été réalisées suite à une extraction manuelle des régions sans utiliser la phase d'extraction automatique de ces régions.

Dans les chapitres suivants nous présentons les solutions que nous avons proposées et étudiées à partir de la chaîne de traitement de Romain et al. Notre première proposition scientifique concerne une nouvelle méthode pour la reconnaissance des polypes basée sur des arbres flous. Notre seconde proposition concerne l'intégration de cette méthode de reconnaissance dans une chaîne de traitement avec l'extraction automatique des régions d'intérêt.

Chapitre 3

Classification des polypes par arbres flous

*Et si nous y mettions un peu de flou ?...
Flou comme la vérité.*

D. Pennac - Monsieur Malaussène

Pour une description plus approfondie des arbres flous le lecteur pourra se reporter à la thèse de Christophe Marsala [68].

Un arbre flou Φ est un algorithme d'apprentissage inductif qui est composée de noeuds, d'arcs et des feuilles organisés en forme d'arbre, voir figure 3.1.

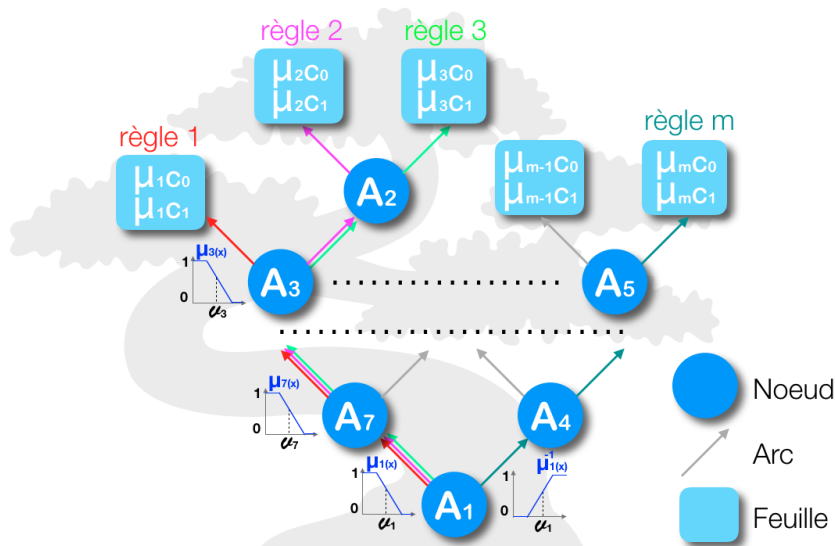


FIGURE 3.1 : Arbre Flou

Dans les noeuds se trouvent les attributs A_i les plus importants. Un attribut A_i est une caractéristique mesurable et observable d'une classe. Dans nos travaux les classes possibles sont "Polype" ou "Non Polype". La position de chaque attribut dans l'arbre est décidée lors de la construction de ce dernier. Dans les arcs se trouvent les

fonctions d'appartenance $\mu_{i(x)}$ qui définissent les points de coupure v_i des attributs A_i . Ils sont calculés lors de la construction de l'arbre flou et sont déterminés à partir des valeurs des attributs A_i .

Dans les feuilles se trouvent des degrés d'appartenance μ_{mC_k} à chaque classe c_k . Ils sont calculés lors de la construction de l'arbre flou. La sortie d'un arbre flou indique le degré d'appartenance à une classe.

Nous avons choisi les arbres flous comme classifieur, pour plusieurs raisons :

- Les arbres flous nous permettent d'accroître la **robustesse du classifieur** ce qui est utile pour gérer la présence d'exemples contradictoires dans une base d'apprentissage [68]. Un haut niveau de robustesse est souhaité pour lever l'incertitude sur certaines images provenant par exemple d'une base où certaines images ont été mal annotées. L'indication d'un degré d'appartenance plutôt qu'une décision tranchée peut augmenter les performances du classifieur en indiquant au gastro-entérologue le "peut-être" là où avec d'autres classifieurs il n'y a que le "oui" et le "non".
- Les arbres flous présentent l'avantage de tenir compte du manque de connaissance des attributs ou des descripteurs [68].
- Les arbres flous ont un grand degré d'intégration sur un système sur puce, comme cela a été démontré dans [18], ce qui répond bien à notre ambition d'intégration dans une **CE**.

Dans cette section nous allons analyser la capacité des arbres flous pour la reconnaissance des polypes. Les images sur lesquelles ces arbres flous sont utilisés sont des régions d'intérêt (**ROI**) provenant d'images de coloscopie. Ces régions sont des régions de l'image où il existe une suspicion de présence d'un polype.

Pour réaliser cette analyse, dans un premier temps nous avons extrait les **ROI** à la main suivant la méthode décrite dans la section 3.1.

Notre analyse est réalisée en quatre étapes : une étape d'extraction des ROI, une étape d'observation, une étape d'apprentissage et une étape de test comme cela est représenté sur la figure 3.2.

- Les **ROI** sont extraites à partir de la base de données d'images de coloscopies ASU-Mayo, présentée dans le tableau 2.1 au chapitre 2. Cette base est annotée et possède une vérité terrain.
- Dans l'étape d'observation des descripteurs de texture et de luminosité pour chaque **ROI** sont extraits. Ces descripteurs sont utilisés pour créer une base d'apprentissage et une base de test.

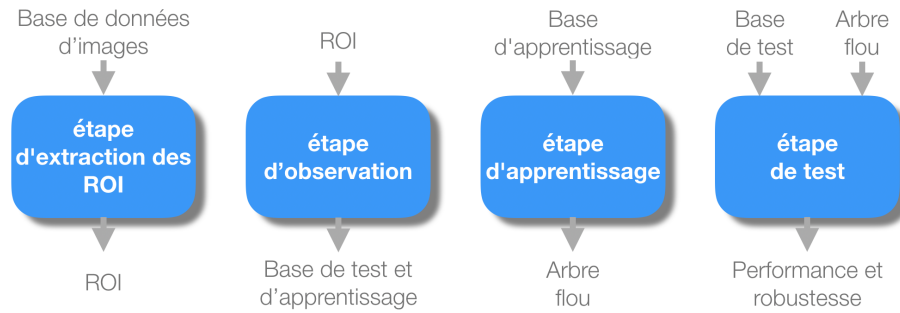


FIGURE 3.2 : Analyse des arbres flous

- Dans l'étape d'apprentissage les arbres flous sont construits à partir des exemples de la base d'apprentissage.
- Dans l'étape de test nous utilisons les arbres flous pour classer des exemples de la base de test et nous mesurons les performances du classifieur réalisé et sa robustesse.

Nous allons décrire en détail chaque étape réalisée pour cette analyse.

3.1 Extraction des ROI

Pour entraîner les arbres flous nous avons utilisé les 20 films de la base de données ASU-Mayo [89] présenté dans le tableau 2.1 au chapitre 2. Dans cette base de données les images ont été annotées par des experts. Chaque image appartient à une des deux classes possible : image avec polype ou image sans polype. Cette annotation est réalisée sous forme d'un masque binaire associé à l'image. Une zone blanche dans le masque binaire indique la présence d'un polype et sa localisation, comme cela est visible sur la figure 3.3.

Nous avons utilisé le masque pour extraire des ROI et chaque ROI a été associé à une classe. Pour avoir une bonne performance, il est conseillé de n'utiliser que deux classes par arbre flou [68]. Nous avons arbitrairement choisi que la classe "1" serait celle qui correspond à la présence d'un polype et la classe "0" l'absence de polype. L'extraction a été réalisée comme suit :

- Pour la classe "1", la ROI a été extraite à partir du masque qui indique la localisation du polype. Une seule ROI est extraite pour chaque image contenant un polype. La taille de la ROI est la même que celle du masque, elle est inférieure à la taille $W \times H$ de l'image.
- Pour la classe "0", plusieurs ROI sont extraites aléatoirement avec des tailles qui varient entre la taille minimale et la taille maximale que peut avoir une ROI de classe "1". Pour une image de taille $W \times H$, nous avons pris $W/80$ comme

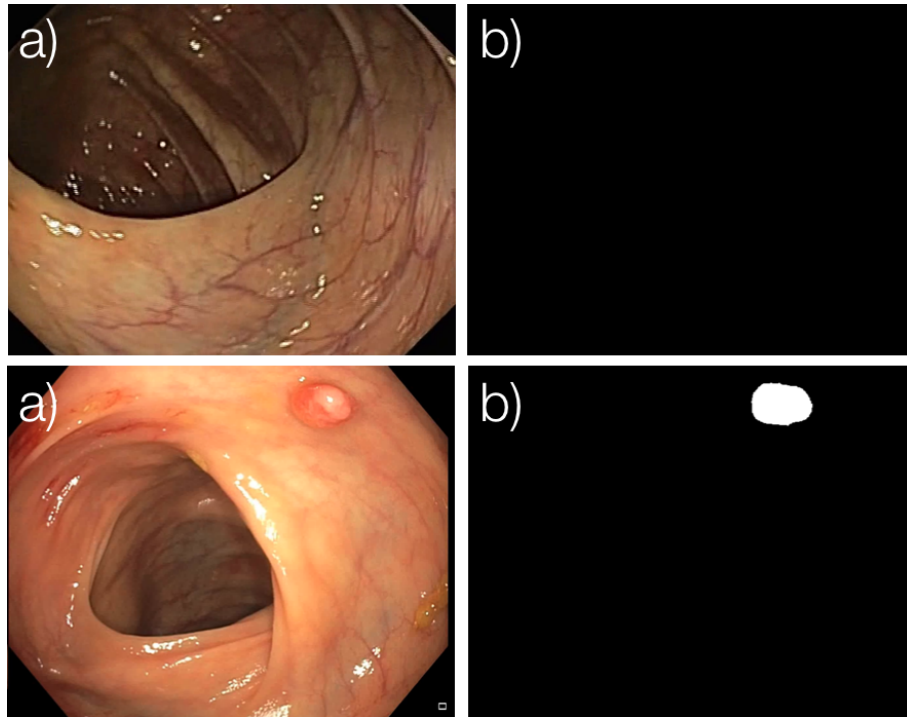


FIGURE 3.3 : a) image , b) masque binaire associé à l'image.

taille minimale et $W/5$ comme largeur maximale d'une ROI, la hauteur d'une ROI est toujours inférieur à sa largeur. Avec cette stratégie il y aura des ROI dans la base d'apprentissage avec la même taille mais de différentes classes.

À l'aide de toutes les ROI extraites nous avons construit une nouvelle base de données de ROI, contenant 131038 ROI dont 3856 ROI de classe "1".

3.2 Description des ROI

Pour chaque ROI nous avons extrait des descripteurs de texture et luminosité pour construire cette base d'apprentissage. Nous avons choisi d'extraire des descripteurs de texture et luminosité parce que ces descripteurs sont discriminants pour la reconnaissance des polypes, voir la section 2.1. De plus la méthode utilisée peut-être intégrée en temps réel dans un système sur puce [17].

Cette méthode utilise le calcul de matrices de co-occurrence [49]. C'est une méthode statistique pour caractériser la périodicité et la directivité des textures dans une image. La matrice de co-occurrence mesure la probabilité pour qu'un couple de niveaux de gris, vérifiant une loi spatiale donnée, apparaisse dans l'image. Le niveau de gris d'un pixel d'une image dépend fortement des niveaux de gris des pixels voisins. La légitimité de la démarche provient d'une modélisation Markovienne des textures : la connaissance de toute l'image autour d'un pixel donné est pratiquement

équivalente à la connaissance du seul voisinage autour de ce pixel. Pour une image I de taille $W * H$, cette matrice M pour la direction horizontale (0°) est calculée à partir de l'algorithme 1, d étant la profondeur du voisinage autour du pixel. Cette matrice peut être calculée dans d'autres directions (45° , 90° et 135°).

Algorithm 1 Calcul de la matrice de co-occurrence

```

1:  $M(0 : 255, 0 : 255) = 0$ 
2: for each column  $i$  from  $d$  to  $W - d$  do
3:   for each row  $j$  from  $d$  to  $H - d$  do
4:      $M(I_{(i,j)}, I_{(i,j+d)}) = M(I_{(i,j)}, I_{(i,j+d)}) + 1$ 
5:      $M(I_{(i,j)}, I_{(i,j-d)}) = M(I_{(i,j)}, I_{(i,j-d)}) + 1$ 
6:   end for
7: end for
  
```

Dans Haralick et al. [49], il a été défini des descripteurs à partir de cette matrice pour résumer les diverses propriétés de la texture, dans le but d'utiliser cette information afin de réaliser une classification.

Nous avons calculé la matrice de co-occurrence dans la direction horizontale (0°) pour une profondeur $d = 1$ et nous aussi avons calculé l'histogramme de luminosité pour chaque ROI.

Nous avons extrait 4 descripteurs statistiques à partir de l'histogramme de luminosité et 22 descripteurs statistiques proposé par Haralick [49]. 26 descripteurs au total ont été calculés. Les descripteurs sont décrits dans l'annexe A.1.

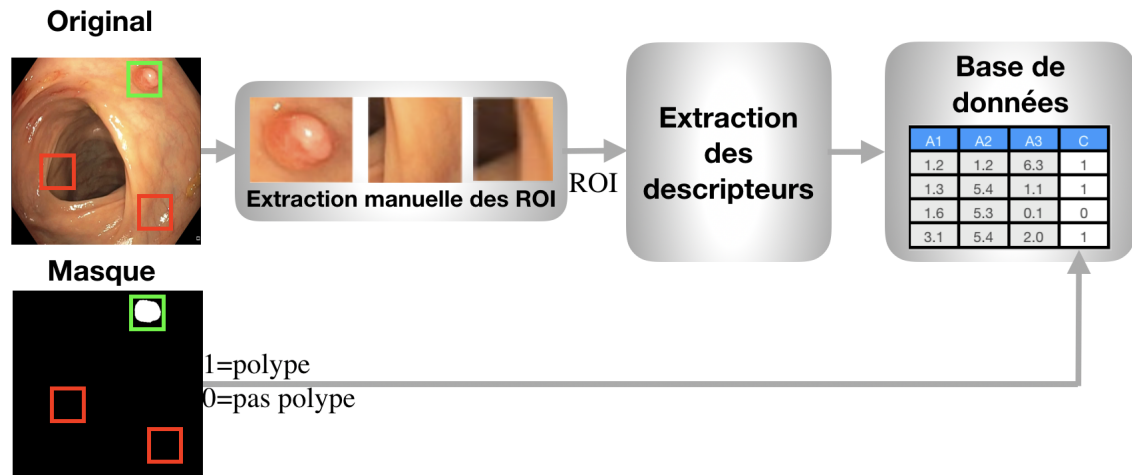


FIGURE 3.4 : Extraction des ROI et calcul des descripteurs.

La figure 3.4 montre la création de la base d'apprentissage en partant des images originales. Cela correspond aux deux premières étapes de notre analyse : l'extraction des ROI et le calcul des descripteurs de chaque ROI.

3.3 Apprentissage des arbres flous

À partir des 131038 ROI extraites. Nous avons distribué aléatoirement 70% soit 91727 de ces ROI pour la base d'apprentissage et 30% soit 39312 pour la base de test.

Nous avons construit 50 bases d'apprentissages équilibrées, constituées de 4628 ROI, soit 2314 ROI pour chaque classe, prises aléatoirement parmi les 91727 de la base d'apprentissage. Nous avons construit une seule base de test composée de 6082 ROI dont 1542 ROI appartenant à la classe "1" prises aléatoirement de la base d'évaluation.

Chaque base d'apprentissage, $\varepsilon\{A_1, A_2, \dots, A_{26}, C_{0,1}\}$, correspond à un tableau de 27 colonnes correspondant aux 26 attributs, ou descripteurs, (A_1 à A_{26}) et à leur classe d'appartenance $C_{0,1}$ dont chaque ligne correspond à un exemple ε_i .

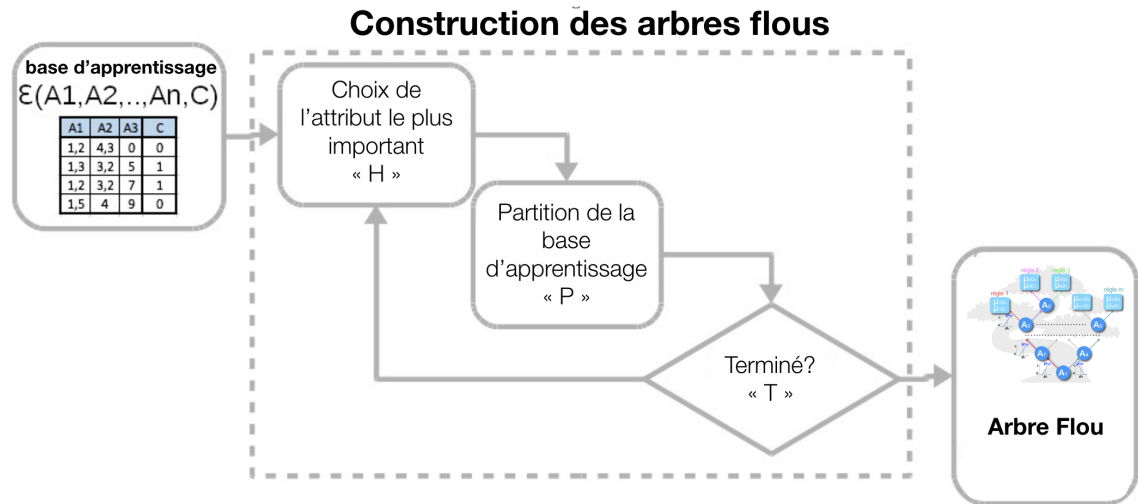


FIGURE 3.5 : Algorithme de construction d'un arbre flou.

Pour construire un arbre flou $\Phi(\varepsilon_{\{A_i, C_k\}}, H, P, T)$ à partir d'une base d'apprentissage nous avons utilisé la méthode décrite en détail dans la thèse de C. Marsala [68], visible sur la figure 3.5. Cette méthode est la suivante :

- L'attribut A_i le plus important est choisi à partir d'une mesure d'information H ; Avant de calculer la mesure d'information H , pour chaque itération chaque attribut A_i est d'abord discrétisé. Il s'agit ici de trouver les fonctions d'appartenance $\mu_{i(x)}$ et leur points de coupure v_i .
- La partition de la base d'apprentissage en deux sous-bases avec le critère de partition P est réalisée ;
- A partir des sous-bases créées une nouvelle discrétisation des attributs est faite.

- Ce processus est itéré jusqu'à ce que chaque sous-base soit sous la forme décrite pour le critère d'arrêt T .

Les méthodes pour construire un arbre flou se différencient entre elles par la mesure d'information H , le critère de partition P et le critère d'arrêt T qu'elle utilisent. Dans notre cas, nous avons utilisé la méthode Salammbô introduite par C. Marsala [68] et nous décrivons dans les parties qui suivent comment sont établis ses différents paramètres H , P et T .

3.3.1 Discrétisation des attributs

La valeur numérique unique d'un attribut est trop précise pour l'apprentissage. Deux valeurs numériques proches portent souvent des significations similaires, mais peuvent être considérées comme des attributs différents dans la construction d'un arbre parce que les algorithmes qui sont utilisés ont été conçus pour manipuler des données symboliques [33]. Ces valeurs sont trop précises et rendent impossible la généralisation, qui est la capacité à classer des données inconnues n'appartenant pas à la base d'apprentissage [68]. Pour permettre cette généralisation et faire perdre la spécialisation d'une valeur numérique d'un attribut, un processus de discrétisation est nécessaire afin de regrouper les valeurs numériques ayant des significations similaires.

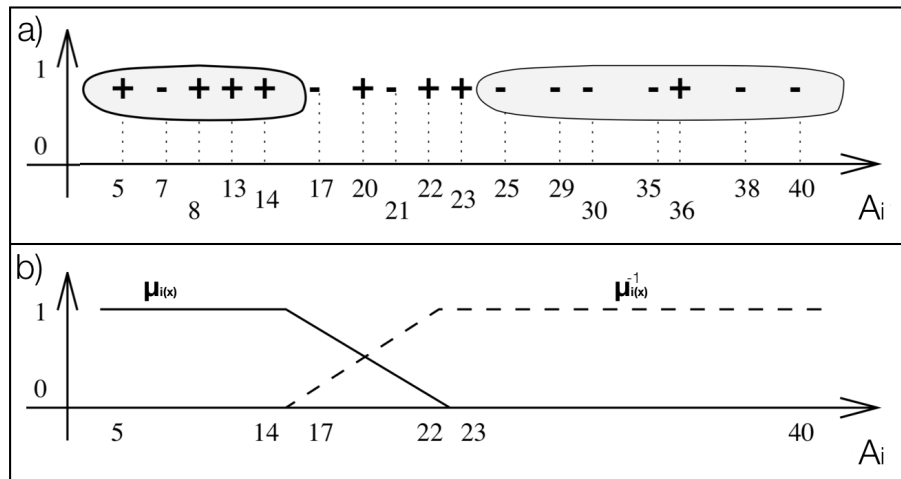


FIGURE 3.6 : a) Valeurs de l'attribut A_i , b) fonction d'appartenance de l'attribut A_i discrétisé

Il existe plusieurs méthodes pour faire cette discrétisation des valeurs des attributs : les méthodes de Randani, Janikow, Yuan et Shaw, CART, Irani, Van Merskt, Funister, etc ... [68]. Nous avons utilisé la méthode proposée pour C. Marsala et F. Bouchon [69] appelée [Generalized Fuzzy Partition using Mathematical Morphology \(FPMM\)](#).

L'algorithme **FPMM** permet de créer un opérateur de filtrage des sous-ensembles flous, voir figure 3.6.

La fonctions d'appartenance $\mu_{i(x)}$ est le résultat de la généralisation des valeurs de l'attribut A_i de la base ou sous base ε .

À partir de cette discrétisation les fonctions d'appartenance $\mu_{i(x)}$ de chaque attribut A_i sont obtenues.

Le point de coupure v_i de la fonction d'appartenance $\mu_{i(x)}$ correspond aux valeurs de l'attribut A_i dont le degré d'appartenance est égal à $\frac{1}{2}$, voir équation 3.1.

$$\mu_i(v_i) = \frac{1}{2} \quad (3.1)$$

3.3.2 Mesure d'information H

La mesure d'information s'utilise pour discriminer l'attribut le plus important d'un ensemble d'attributs. C'est à dire celui dont l'information est la plus importante pour prendre la décision. La méthode la plus utilisée pour faire la mesure d'information est l'entropie de Shannon[84], cette entropie est utilisé dans les méthodes de ID3 [75] et CART [93], mais elles ne prennent pas en compte les données floues. Pour traiter des données floues, la méthode la plus utilisée est l'entropie en étoile. L'entropie en étoile est une extension de l'entropie de Shannon obtenue en substituant la mesure de probabilité classique par une mesure de probabilité d'évènements flous. La mesure de probabilité d'évènements flous généralement utilisée est celle proposée par Zadeh[96], l'entropie en étoile est calculée à l'aide de l'équation 3.2, où " K " est le nombre d'exemples dans la sous-base.

$$H_{z(\mu_{i(x)})} = - \sum_{k=1}^K \mu_{i(k)} P_{(\mu_{i(k)})} \log P_{(\mu_{i(k)})} \quad (3.2)$$

Nous avons utilisé l'entropie en étoile pour mesurer l'information à l'aide de H .

3.3.3 Stratégie de Partition P

Pour réaliser la partition de la base originale en deux sous-bases nous utilisons une partition alpha-couple avec ($\alpha = \frac{1}{2}$). Cette partition est décrite par l'équation 3.3. Chaque sous-base ε sera composée de tous les exemples e_i dont le degré d'appartenance $\mu_i(e_{i(A_i)})$ de l'attribut A_i est plus grand que α .

$$\varepsilon = e_i \in \varepsilon / \mu_i(e_i(A_i)) \geq \alpha \quad (3.3)$$

3.3.4 Critère d'arrêt T

Notre critère d'arrêt dit que la construction est terminée quand tous les exemples d'une sous-base ont la même classe ou quand il y a au moins " n " exemples dans une sous-base. Dans notre cas nous utilisons $n = 2$.

À la fin de cette phase d'apprentissage un arbre flou est obtenu. Il est composé par M règles. Chaque règle m est composée des noeuds, des arcs et des feuilles. Les noeuds A_j correspondent à l'attribut j . Les arcs correspondent à la fonction d'appartenance $\mu_{j(x)}$ et son point de coupure v_j pour chaque attribut A_j . Les feuilles correspondent aux degrés d'appartenance μ_{mC_0} et μ_{mC_1} de la règle m pour chaque classe.

3.4 Test de la classification des ROI avec les arbres flous

Nous avons construit 50 arbres flous à partir des 50 bases d'apprentissage.

Nous avons utilisé deux méthodes de classification pour indiquer la classe de chaque ROI(w_1, w_2, \dots, w_{26}) extraite. Une classification binaire et une classification floue [68].

La classification binaire utilise la méthode du Modus Ponem classique et la sortie du classifieur est binaire, "0" pour l'absence de polype ou "1" pour la présence de polype.

Dans la méthode du Modus Ponem classique l'arbre flou est traité comme une arbre binaire, en commençant dès le premier noeud. Le degré d'appartenance de la valeur observée w_j est comparé avec le degré d'appartenance du point de coupure dans l'arc v_j pour aller d'un noeud vers le suivant, si $\mu_{j(w_j)} > \mu_{j(v_j)}$ le noeud suivant est considéré. Cette comparaison est réalisée pour chaque attribut j de la règle m jusqu'à la feuille μ_{mC_k} , si μ_{mC_1} est plus grand que μ_{mC_0} la classe de la feuille est "1", elle est "0" dans le cas contraire.

La classification floue utilise la méthode Modus Ponem généralisée et à la sortie un degré d'appartenance est associé à chaque classe μ_0 et μ_1 .

La méthode Modus Ponem généralisée est la suivante :

1. un degré de satisfiabilité, $Fded_{m(c_k)}$, est calculé pour chaque classe $c_k = \{0, 1\}$ et pour chaque règle m . Pour calculer ce degré un opérateur de norme \top est utilisé parmi ceux définis dans le tableau 3.1. Les valeurs de l'objet observé w_j et leurs fonctions d'appartenance $\mu_{j(w_j)}$, correspondant à chaque attribut j de la règle m , sont multipliés par un facteur connu comme étant la probabilité

conditionnelle de Zadeh $\mu_{mC_{c_k}} = P_{(c_k/(v_1, v_2, \dots, v_j))}^*$.

$$Fded_{m(1)} = \top_{j=1, \dots, J} \mu_j(w_j) * \mu_{mC_1} \quad (3.4)$$

$$Fded_{m(0)} = \top_{j=1, \dots, J} \mu_j(w_j) * \mu_{mC_0} \quad (3.5)$$

TABLE 3.1 : Opérateurs de norme \top et conorme \perp .

Name	norme \top	conorme \perp
Zadeh	$\min(x, y)$	$\max(x, y)$
Lukasiewicz	$\max(x + y - 1, 0)$	$\min(x + y, 1)$
Boole	xy	$x + y - xy$
Hamacher	$xy / (x + y - xy)$	$(x + y - 2xy) / (1 - xy)$
Einstein	$xy / (2 - x - y - xy)$	$xy / (1 + xy)$

2. Un nouveau degré d'appartenance μ_{c_k} est calculé en agrégeant tous les degrés de satisfiabilité de chaque règle m . L'agrégation est faite avec un opérateur de conorme \perp parmi ceux définis dans le tableau 3.1.

$$\mu_0 = \perp_{m=1, 2, \dots, M} Fded_{m(0)} = \mu_{non_polyp} \quad (3.6)$$

$$\mu_1 = \perp_{m=1, 2, \dots, M} Fded_{m(1)} = \mu_{polyp} \quad (3.7)$$

Pour la classification floue nous avons testé les cinq opérateurs de norme et de conorme triangulaire définis dans le tableau 3.1.

Pour évaluer les performances nous avons calculé les métriques suivantes :

$$Sensibilité = VP / (VP + FN) \quad (3.8)$$

$$Spécificité = VN / (VN + FP) \quad (3.9)$$

où :

- **Vrai Positif (VP)** : C'est une ROI avec un polype qui a été classé comme "1" ;
- **Faux Positif (FP)** : C'est une ROI sans polype qui a été classé comme "1" ;
- **Vrai Négatif (VN)** : C'est une ROI sans polype qui a été classé comme "0" ;
- **Faux Négatif (FN)** : C'est une ROI avec un polype qui a été classé comme "0".

La sensibilité mesure la capacité à reconnaître une [ROI](#) contenant un polype et la spécificité mesure la capacité de ne pas commettre un erreur de classification, soit la capacité de n'est pas avoir des faux positifs.

Les résultats que nous avons obtenus sont illustrés dans la figure 3.7.

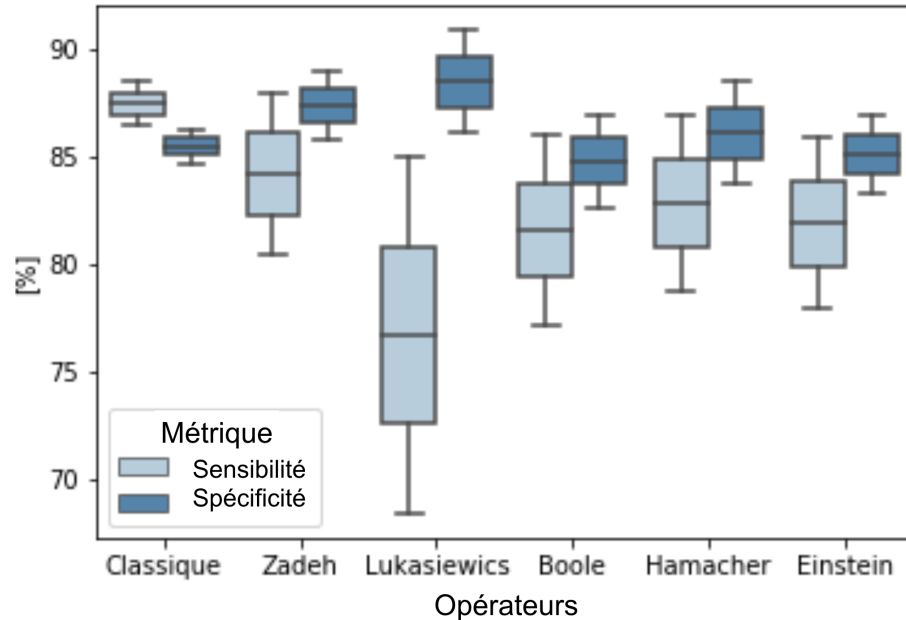


FIGURE 3.7 : Résultats de classification pour 50 Arbres Flous.

À partir de l'analyse de ces résultats, nous pouvons voir que la méthode de classification binaire, dite classique, possède une bonne stabilité avec une variation de 1%. Elle possède sensibilité de 88% et une spécificité de 86%. Ces performances sont bonnes et dépassent celles des tests de diagnostics précoces, tel que ceux liés aux dépistages, actuellement utilisés. Les méthodes de classification floues ont de bonnes capacités de reconnaissance, par contre elles sont moins stables, le meilleur des opérateurs pour notre cas est celui de Zadeh avec une variation de 3%.

La classification binaire montre une meilleure stabilité pour les 50 arbres. Ceci est normal parce que les arbres flous nous permettent d'accroître la robustesse du classifieur. Cet effet obtenu grâce à la discrétisation des attributs faites pendant la construction de l'arbre. L'algorithme utilisé, [FPMM](#), permet que des exemples contradictoires soient ignorés, comme cela est visible dans la figure 3.6.

Par contre pour la classification floue, la décision se fait à partir des valeurs des noeuds, arcs et feuilles de l'arbre flou et de l'**opérateur d'agrégation** utilisé. La stabilité dépend du type d'opérateur utilisé. Dans notre cas les opérateurs de Zadeh montre un comportement plus stable que les autres opérateurs. Malgré cette instabilité, l'avantage de la classification floue est qu'elle calcule un degré d'appartenance aux classes entre 0 et 1, ce qui est une information exploitable pour agréger plusieurs classifieurs spécialisés pour différentes classes comme différents types de polypes.

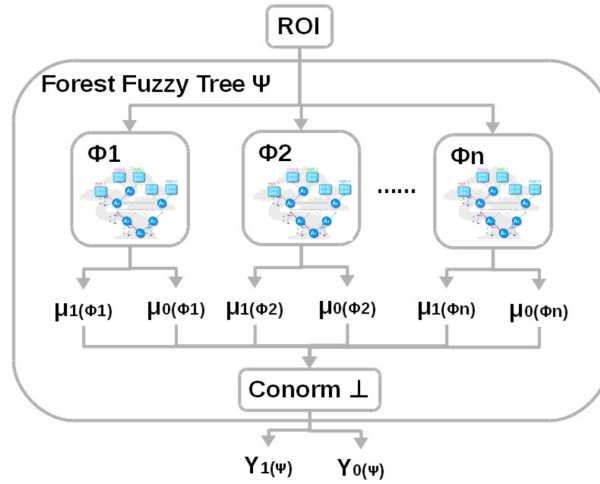


FIGURE 3.8 : Forêt Flou

Pour améliorer la performance de reconnaissance des arbres flous nous proposons de construire une **forêt d'arbres flous** $\Psi\{\Phi_1, \Phi_2, \dots, \Phi_n\}$ en agrégeant n arbres flous Φ_n .

3.5 Classification des ROI avec une forêt d'arbres flous

La forêt $\Psi\{\Phi_1, \Phi_2, \dots, \Phi_n\}$ que nous avons construite est modélisée par le diagramme de la figure 3.8.

L'idée est de calculer un nouveau degré d'appartenance à partir de l'agrégation des degrés d'appartenance des n arbres flous $\forall_{i=1,2,\dots,n} \{\mu_{polyp}(\Phi_i), \mu_{n_polyp}(\Phi_i)\}$ en utilisant un opérateur de conorme \perp .

$$\Upsilon_0 = \perp_{i=1,2,\dots,n} \{\mu_0(\Phi_i)\} = \Upsilon_{n_polyp} \quad (3.10)$$

$$\Upsilon_1 = \perp_{i=1,2,\dots,n} \{\mu_1(\Phi_i)\} = \Upsilon_{polyp} \quad (3.11)$$

Dans notre cas, pour la classification floue nous avons utilisé comme conorme les opérateurs de Zadeh.

Si le degré d'appartenance de la classe "1", Υ_1 , est supérieur au degré d'appartenance de la classe "0", Υ_0 , la **ROI** est classée comme "1" sinon elle est classée comme "0".

Pour la classification binaire nous proposons un vote majoritaire comme conorme. Si la plupart des arbres classent la **ROI** comme "1", elle est finalement classée comme "1" sinon elle est classée comme "0".

Les résultats obtenus sont visibles sur la figure 3.9. La performance de détection s'est clairement améliorée pour les deux méthodes, dans le cas de classification bi-

naire la capacité de reconnaissance a été augmentée jusqu'à une sensibilité de 94% et une spécificité de 89%. La performance de la méthode de classification floue a augmenté jusqu'à une sensibilité de 90% et une spécificité de 85%.

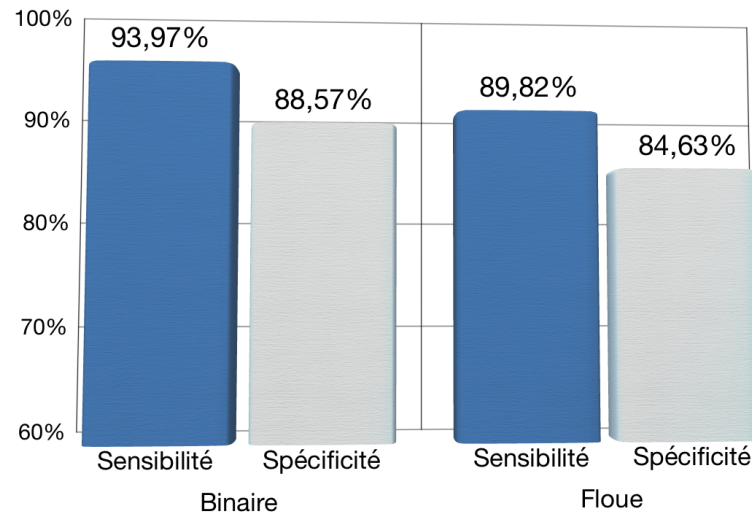


FIGURE 3.9 : Résultats de la classification avec la forêt de 50 arbres flous.

3.6 Discussion

Dans cette section nous avons évalué la capacité des arbres flous à reconnaître les polypes sur des ROI extraites à la main à partir d'images provenant de coloscopie.

Les tests ont montré que les arbres flous obtiennent des performances de classification entre 85% et 91% mais avec des variations pouvant être importantes dans l'écart type en fonction de l'opérateur utilisé. La classification binaire ayant un écart type limité démontre une meilleure stabilité que les méthodes floues.

L'utilisation d'une forêt d'arbres flous a permis d'augmenter généralement la performance de reconnaissance pour toutes les méthodes de classification permettant d'atteindre des taux de classification de 93% avec une spécificité de 91% pour la méthode binaire, et un taux de classification de 90% avec une spécificité de 85% pour la méthode floue.

Ces résultats nous indiquent que :

- Les descripteurs des texture et luminosité sont des descripteurs pertinents et discriminants dans la reconnaissance des polypes ;
- Les arbres flous ont une stabilité élevée dans la reconnaissance des polypes ;
- La forêt floue permet d'améliorer globalement les performances des arbres flous ;

Dans notre cas d'étude les arbres utilisant la classification binaire ont une meilleure performance que les méthodes de classifications floue. Ceci s'explique par la répartition des exemples de la base de données en seulement deux classes : polypes et non polypes. Une étude serait à mener en perspective à ces travaux avec une labelisation à plus de 2 classes pour analyser plus en détail la pertinence d'une classification floue.

Au vue de nos résultats, la méthode que nous allons adopter pour la suite de notre étude est la classification binaire.

Dans cette première étude nous n'avons pas considérée l'extraction automatique des [ROI](#). Cette extraction va avoir un impact important sur les performances de la chaîne puisqu'il s'agit ici de fournir des données extraites automatiquement au classifieur, données qui doivent avoir la même qualité que celles qui ont permis de valider les arbres flous.

Dans la section suivante nous allons évaluer une méthode d'extraction automatique des [ROI](#) et déterminer son impact dans une chaîne globale à partir de l'image jusqu'à la classification en utilisant des arbres flous comme classifieur.

Chapitre 4

Extraction automatique des ROI

Instruit par la mésaventure des miens, j'avais appris à me méfier des évidences. Lorsque tout le monde s'agglutine autour d'une même opinion, je m'enfuis : la vérité est sûrement ailleurs.

A. Maalouf - Léon l'africain

Nous désirons évaluer la méthode d'extraction automatique des ROI et déterminer son impact dans une chaîne globale à partir de l'image jusqu'à la classification en utilisant les forêts flous. Pour cela nous proposons de mettre en œuvre cette chaîne en trois étapes principales. Ces étapes, qui sont visibles sur la figure 4.1, sont :

- **L'extraction des ROI** : dans cette étape des zones de l'image sont identifiées comme pouvant contenir des polypes, elles deviennent des ROI et vont être analysées ;
- **La description des ROI** : dans cette étape pour chaque ROI des caractéristiques sont extraites, et des descripteurs sont calculés pour les représenter ;
- **La classification des ROI** : dans cette étape à l'aide des descripteurs calculés précédemment, un classifieur, préalablement entraîné, classe chaque ROI soit comme contenant un polype (classe "1"), soit comme n'en contenant pas (classe "0").



FIGURE 4.1 : Diagramme d'une chaîne de traitement pour la reconnaissance des polypes.

Dans [53] les auteurs remarquent que l'étape la plus critique pour la détection automatique des polypes est le prétraitement, qui dans notre cas correspond à l'extraction des ROI.

La méthode d'extraction automatique des ROI que nous proposons s'inspire de la pratique des gastroentérologues lors d'une coloscopie :

1. Dans un premier temps le physicien réalise une recherche circulaire ou linéaire dans l'image avec l'objectif de sélectionner des ROI dans la paroi intestinale en cherchant des formes qui ressemblent aux polypes, cercles ou ellipses.
2. Ensuite il analyse les ROI et grâce à son expertise détermine la présence possible d'un polype.

La lecture d'un certains nombre de travaux [54, 53, 78, 67], confirme que les polypes sont souvent des protubérances avec des formes circulaires et/ou elliptiques et peuvent être repérés par leurs courbures locales.

Pour extraire des ROI, nous allons balayer toute l'image en cherchant des objets avec une forme circulaire et/ou elliptique. Pour cela nous utilisons la transformée de Hough pour détecter ces objets circulaires/elliptiques. Cette transformée est connue pour son efficacité dans la détection des cercles et son haut niveau d'intégration [100], [83], [52], [58].

Notre méthode pour extraire des ROI est la suivante :

- Dans un premier temps, l'image est convertie de l'espace des couleurs RGB à une image de luminance en niveau de niveau de gris. Nous avons utilisé l'équation 4.1 pour faire cette conversion. Nous effectuons cette transformation parce que nous désirons ne travailler que sur la texture de l'image pour détecter les cercles et que la luminance préserve l'information de texture originale. Cela réduit de plus les calculs puisque les images n'utilisent plus qu'un seul canal contre trois pour une image RGB. Cette réduction réduit les contraintes d'intégration, ce que nous cherchons à faire dans le cadre de nos travaux.

$$Y = 16 + \frac{66}{256}R + \frac{129}{256}G + \frac{25}{256}B \quad (4.1)$$

- Un filtre de bruit utilisant un filtre médian 3x3 est ensuite appliqué.
- Les contours sont ensuite identifiés à l'aide du filtre de Canny [31].
- Finalement, la transformée de Hough est utilisée pour trouver des contours circulaires/elliptiques. Une ROI sera extraite sur toute région de l'image où un cercle ou une ellipse sera trouvée, voir figure 4.2.

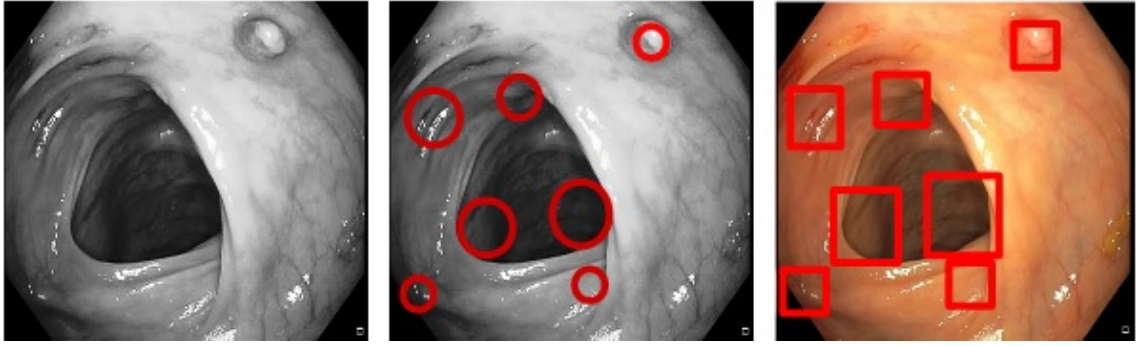


FIGURE 4.2 : Étape d'extraction des ROI.

4.1 Experimentations

TABLE 4.1 : Détails de partage de la base de données des images de coloscopie ASU-Mayo [89].

Film	Base d'apprentissage			
	Total (100%) Images (avec polype)	Résolution	Apprentissage (70%) Images (avec polype)	Test (30%) Images (avec polype)
1	682(0)	712x480	477(0)	205(0)
2	838(0)	712x480	586(0)	252(0)
3	769(0)	712x480	538(0)	231(0)
4	712(0)	712x480	498(0)	214(0)
5	1843(0)	712x480	1290(0)	553(0)
6	1925(0)	712x480	1347(0)	578(0)
7	1550(0)	712x480	1085(0)	465(0)
8	1740(0)	712x480	1218(0)	522(0)
9	1802(0)	712x480	1261(0)	541(0)
10	1639(0)	712x480	1147(0)	491(0)
11	324(245)	1920x1080	226(171)	98(74)
12	910(910)	1920x1080	637(637)	273(273)
13	519(374)	1920x1080	363(261)	156(113)
14	501(391)	856x480	350(273)	151(118)
15	1200(1106)	856x480	840(774)	360(332)
16	339(209)	1920x1080	237(146)	102(63)
17	418(234)	856x480	292(163)	126(71)
18	259(189)	1920x1080	181(132)	78(57)
19	616(235)	1920x1080	431(164)	185(71)
20	410(385)	856x480	287(269)	123(116)
Total	18996(4278)		13291(2990)	5704(1288)

Pour réaliser nos expérimentations nous avons partagé aléatoirement la base de données ASU-Mayo [89] de 18996 images de coloscopies en deux parties :

- une base d'apprentissage constituée de 70% de la base, soit 13291 images ;
- une base de test constituée de 30% de la base, soit 5704 images.

La distribution est détaillée dans le tableau 4.1.

À partir des images de la base d'apprentissage, 10 bases équilibrées constituées de 4628 **ROI**, soit 2314 **ROI** pour chaque classe, ont été construites. Nous avons utilisé ces bases pour entraîner une forêt floue composé de 10 arbres flous en utilisant la méthode de construction décrite dans le chapitre 3.3 avec des **ROI** extraites manuellement.

Pour analyser l'impact de notre méthode automatique d'extraction des **ROI**, nous l'avons évalué à travers les performances de détection de l'ensemble de la chaîne : extraction des **ROI** et classification.

Ces mesures de performances ont été réalisées sur les images de la base de test.

Pour les performances au **niveau des ROI**, c'est le classifieur qui définit la classe, soit classe "1" pour présence de polype ou classe "0" pour l'absence de polype.

Pour les performances au **niveau image**, la décision est prise comme suit : une image est classée classe "1" si elle contient au moins une **ROI** de classe "1", si non elle est classée classe "0".

Les métriques que nous avons utilisées pour cette évaluation sont la sensibilité et la spécificité. Les résultats sont visibles dans le tableau 4.2.

TABLE 4.2 : Performances de détection de polypes

Films	Résolution [pixels]	Niveau ROI		Niveau Image	
		Sens.	Spec.	Sens.	Spec.
6	1920x1080	37,7%	93,9%	62,9%	49,0%
4	856x480	17,4%	96,3%	40,7%	54,8%
10	712x480	-	95,7%	-	76,2%
20	all	29,9%	95,2%	53,2%	73%

Les résultats montrent une sensibilité de seulement 30% au niveau **ROI** et de 53% au niveau image avec une spécificité de 73%. Cela contraste fortement avec les résultats obtenus précédemment et présentés dans le chapitre 3 où la sensibilité pouvait atteindre jusqu'à 93% et la spécificité jusqu'à 90%. La réduction de sensibilité est de 60% et cela montre l'impact de l'extraction automatique des **ROI** sur la performance globale de classification.

Nous avons étudié la raison de cette perte de performance du classifieur. Ce que nous avons observé, c'est que quand les **ROI** ont été extraites manuellement, le polype est centré et entièrement contenu dans la **ROI**. Ceci n'est plus vrai dans le cas d'une extraction automatique, les **ROI** ne peuvent contenir qu'une partie du polype et celui-ci peut-être n'importe où dans la **ROI**. La figure 4.3 montre l'extraction des polypes à la main et de façon automatique :

- dans sa partie a) nous pouvons voir l'extraction à la main.

- dans sa partie b) nous pouvons voir l'extraction automatique pour les mêmes images de polypes.

Nous nous apercevons que dans les cas 1 et 3 dans la partie b) les polypes ne sont pas centrés dans la ROI. Pour le cas 2 dans la partie b), le polype est centré mais seulement une partie est extraite, la ROI est plus petite que le polype. Pour le cas 4 dans la partie b), c'est l'inverse la ROI est trop grande par rapport au polype qui est tout de même centré.

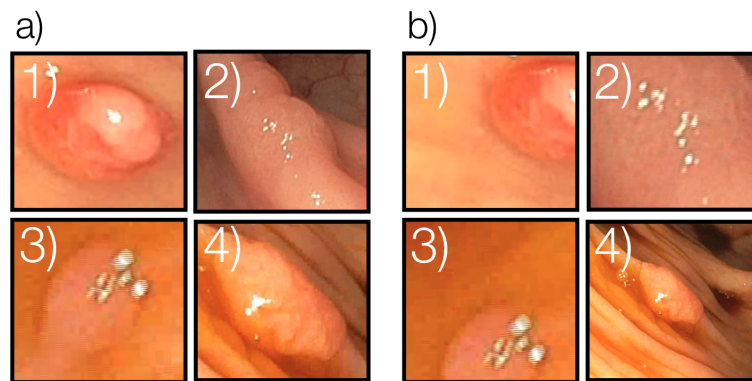


FIGURE 4.3 : a) ROI extraites à la main, b) ROI extraites automatiquement.

Il faut ajouter à cette constatation, l'impact des performances de l'extraction en elle-même qui ne garantit pas que toutes les ROI contenant un polype seront extraites, ce qui était le cas dans une extraction manuelle.

Une analyse détaillée de la performance d'extraction des ROI contenant un polype au moins, nous a permis de constater que :

- 18,7% des polypes ont été extraits dans des ROI qui contiennent au moins 90% de ce polype ;
- 51,2% des polypes ont été extraits dans des ROI qui contiennent au moins 50% de ce polype ;
- 89,9% des polypes ont été extraits dans des ROI qui contiennent au moins 10% de ce polype ;
- 10,1% des polypes n'ont pas été insérés dans une ROI ;

Cette analyse montre que seulement 89,9% de polypes se retrouvent dans une ROI extraite automatiquement, et que parmi ceux-là, seulement 18,7% de polypes apparaissent quasi-entièrement dans une ROI.

En total 309719 ROI ont été extraites, parmi elles 13743 ROI contiennent un polype. Une analyse détaillée des 13743 ROI contenant un polype nous a permis de constater que :

- 11,5% de ces ROI contiennent entre 100 et 90% d'un polype;
- 31,6% de ces ROI contiennent entre 90 et 50% d'un polype;
- 56,9% de ces ROI contiennent entre 50 et 10% d'un polype.

Cette analyse montre que seulement 11,5% des ROI contenant un polype le contiennent quasi-entièrement.

Ceci est la principale raison pour laquelle le modèle de classifieur, entraîné avec des ROI extraites à la main, obtient des performances faibles avec des ROI extraites automatiquement.

Cet effet est visible dans la performance de sensibilité au niveau ROI et au niveau image puisque notre classifieur a été entraîné pour reconnaître des ROI contenant un polype complet et centré.

Pour évaluer l'impact de l'extraction automatique des ROI nous avons mesuré la sensibilité pour chaque cas, les résultats sont visible sur la figure 4.4.

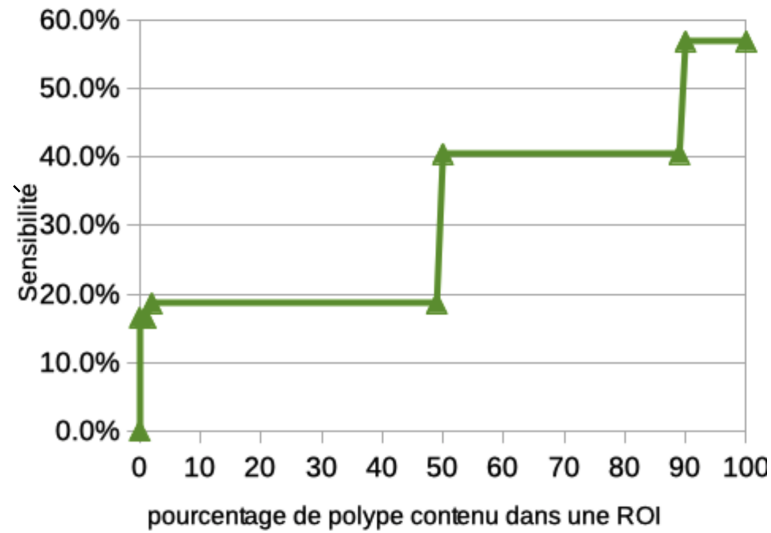


FIGURE 4.4 : Sensibilité de classification selon le pourcentage de polype contenu dans une ROI.

Nous constatons que plus le pourcentage de polype dans une ROI est grand plus la sensibilité du classifieur est grande, ce qui semble cohérent. Il nous faut donc augmenter le nombre des ROI contenant au moins 50% d'un polype.

Au niveau de l'image, la règle initiale utilisée est simple mais se révèle inefficace. Nous avons calculé la spécificité liée à cette règle, elle est égale à la probabilité que dans une image sans polype, toutes les n ROI extraites soient classées comme classe "0", cela correspond à une probabilité binomiale représentée par :

$$Spécificité_{Niveau Image} = (P_{(class_{ROI}=0)})^n = (Spécificité_{Niveau ROI})^n$$

Dans nos expériences le nombre des ROI dans une image sans polype peut varier de 5 à 13, la spécificité peut varier dans ce cas théoriquement de 78% à 53%. Ce que nous constatons puisque nos mesures indiquent une spécificité de 73%. La règle utilisée pour agréger les ROI et donner une décision au niveau image doit être modifiée pour augmenter les performances de la chaîne.

Pour dépasser ces limites, nous avons conçu une nouvelle approche pour la détection automatique des polypes, ce qui est l'objet de la partie suivante.

4.2 Extraction et suivi des ROI pour la détection des polypes

Cette nouvelle méthode est inspirée de la méthodologie psycho-visuelle des gastro-entérologues lors d'une coloscopie.

1. Dans un premier temps l'expert cherche des formes circulaire dans l'image avec l'objectif d'identifier si la paroi intestinale présente une anomalie : un polype.
2. Ensuite l'expert une fois sélectionné la partie de l'image à analyser décide grâce a son expertise si l'anomalie est un polype.
3. Pour prendre cette décision l'expert analyse plus qu'une image. Il observe sur une séquence d'images les caractéristiques de l'anomalie, taille, texture, position etc... . Cette séquence d'images varie en fonction des allers et retours que l'expert réalise avec l'endoscope afin d'explorer la partie d'intestin contenant une possible anomalie.

Il apparaît ici que le gastro-entérologue travaille directement sur un flux vidéo avec une analyse dynamique des images. Une ROI peut être sélectionnée et suivie sur plusieurs images successives avant de prendre une décision.



FIGURE 4.5 : Chaîne de traitement pour l'extraction et suivi des ROI pour la détection automatique des polypes.

A l'aide de ces constatations, nous avons modifié la chaîne de traitement initiale afin de sélectionner des ROI et de les suivre dans les images suivantes. Cette nouvelle chaîne est composée de cinq étapes. Ces étapes, visibles dans la figure 4.5, sont :

1. **L'extraction des ROI** : dans cette étape les ROI sont extraites comme précédemment, tel que cela a été décrit dans le chapitre 3.
2. **Le suivi des ROI** : si dans l'image une ROI a été validé comme appartenant à la classe "1" par le classifieur, une méthode d'estimation du mouvement est utilisée pour déterminer la position de cette ROI dans la prochaine image.
3. **La description des ROI** : dans cette étape des descripteurs de texture et luminosité sont extraits pour décrire chaque ROI. Cette étape a été décrite dans le chapitre 3.
4. **La classification des ROI** : dans cette étape chaque ROI est classée comme contenant un polype ("1") ou pas ("0"). Cette étape a été décrite dans le chapitre 3.
5. **L'agrégation des ROI** : la décision au niveau image est faite à partir de l'agrégation de toutes les ROI de classe "1" dans un flux d'images.

L'étape du suivi des ROI a pour objectif principal de maximiser la surface du polype présent dans une ROI.

A partir d'une image où une ROI est extraite et classée comme classe "1" dans l'image I_{n-1} , nous utilisons sa localisation dans I_{n-1} pour estimer sa localisation dans l'image I_n . Le suivi peut être réalisé d'une image à l'autre et dans ce cas il aura une profondeur temporelle d'une image, mais nous avons aussi expérimenté le suivi jusqu'à une profondeur temporelle de trois images successives. Cette nouvelle ROI est transmise à l'étape de description pour être classée.

L'étape de description des ROI est réalisée après l'étape d'extraction et de suivi des ROI, ces deux étapes sont exécutées en parallèle.

L'étape d'agrégation des ROI détermine la classe de l'image d'une manière plus efficace en exploitant à la fois l'information spatiale et temporelle. Une image est considérée comme appartenant à la classe "1" s'il y a au moins une ROI de classe "1" qui a été aussi classe "1" dans les N dernières images consécutives. Nous avons déterminé de façon empirique que $N = 3$ était une bonne profondeur, comme le montre la figure 4.6. Si une ROI de classe "1" a été classée "0" dans les N derniers images consécutives, il est probable que soit un faux positif, cette nouvelle méthode d'agrégation doit pouvoir diminuer ces faux positifs et en conséquence augmenter le taux de vrai négatifs, la spécificité, au niveau image.

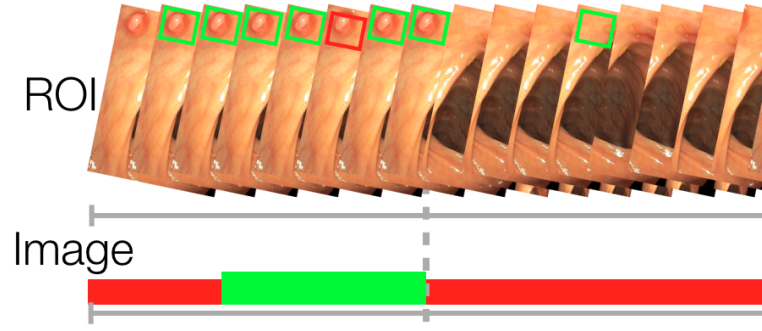


FIGURE 4.6 : Agrégation des ROI.

En plus d'améliorer le centrage des polypes dans les ROI, l'étape de suivi des ROI peut identifier de nouvelles ROI n'ayant pas été identifiés par l'étape d'extraction des ROI.

La méthode pour réaliser l'étape de suivi des ROI est décrite dans la prochaine section.

4.2.1 Suivi des ROI

Pour suivre une ROI nous calculons l'estimation de mouvement en utilisant une **correspondance par blocs**.

Pour ce faire nous réalisons les actions suivantes :

- Chaque ROI validée comme classe "1" dans la dernière image I_{n-1} est considérée comme un **bloc** $B_{p,q}$ de taille $P * Q$;
- Pour tous les pixels dans le bloc $B_{p,q}$, un seul **vecteur de mouvement** est calculé ;
- Le **vecteur des valeurs de l'image d'intensité** $I_{n-1}(B_{p,q})$ d'un bloc $B_{p,q}$ de l'image I_{n-1} est défini par la formule :

$$I_{n-1}(B_{p,q}) = [I_{n-1}(p+1,q), \dots, I_{n-1}(p+P-1,q+Q-1)] \quad (4.2)$$

- Le bloc est déplacé à partir de sa position initiale (p, q) vers $(p-i, q-j)$ par les **vecteurs candidats de mouvement** $\vec{V} = (i, j)$ (voir figure 4.7). Dans notre cas, nous utilisons 8 vecteurs candidats de mouvement et jusqu'à 10 voisinages avec une distance de 2, 4, 6, 8, 10, 12, 15, 20, 25 et 30 pixels respectivement ;
- L'estimation de mouvement est faite en calculant une **mesure de similarité** entre $I_{n-1}(B_{p,q})$ et $I_n(B_{p-i,q-j})$ en utilisant la mesure SAD_{ij} , la **somme des**

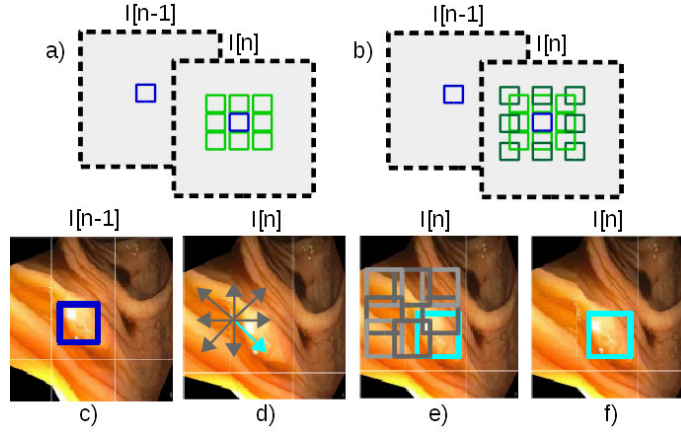


FIGURE 4.7 : a) Correspondance des blocs avec un voisinage d'une distance de 2 pixels. b) Correspondance de blocs avec 2 voisinages d'une distance de 2 et de 4 pixels. c) ROI validé dans l'image I_{n-1} . d) 8 vecteurs candidats de mouvement. e) 8 blocs candidats déplacés par les vecteurs candidats de mouvement. f) Bloc ayant une **somme des valeurs absolues**, SAD_{ij} , minimale.

valeurs absolues des différences définie par l'équation 4.3.

$$SAD_{ij} = \sum_{\forall p,q \in [B*P]} |I_{n-1}(B_{p,q}) - I_n(B_{p-i,q-j})| \quad (4.3)$$

- Le **vecteur candidat de mouvement** \vec{V} avec la mesure SAD_{ij} minimale est considéré comme le **vecteur de mouvement** et le bloc est déplacé de ce vecteur dans l'image I_n ;
- Toutes les ROI classées comme appartenant à la classe "1" dans l'image I_{n-1} vont être les ROI pour l'image I_n déplacées par le **vecteur de mouvement** \vec{V} (voir figure 4.7).

Avec cette technique il est possible de suivre une ROI où un polype a été détecté dans l'image I_{n-1} .

Il est aussi possible d'augmenter la profondeur *depth* de l'estimation de mouvement pour suivre une ROI où un polype a été détecté dans les images I_{n-1} , $I_{n-2}, \dots, I_{n-Depth}$.

L'augmentation de la profondeur *depth* peut améliorer la performance de la chaîne en diminuant le nombre des faux négatifs au niveau image du à des ROI ne contenant pas de polype mais identifiées comme telles et suivies.

4.2.2 Expérimentations

Nous avons évalué la performance de détection au niveau ROI et au niveau image de cette nouvelle méthode sur la base de données précédemment utilisée en calculant

les mêmes métriques : la sensibilité et la spécificité.

Pour analyser l'impact de notre nouvelle agrégation des ROI nous avons utilisé une profondeur d'estimation de mouvement $depth = 0$. Pour analyser l'impact de l'étape de suivi des ROI nous avons utilisé plusieurs valeurs de profondeur $depth = 1, 2, 3$.

Les performances obtenues sont visibles dans le tableau 4.3.

TABLE 4.3 : Performances de détection de polypes

Depth	Niveau ROI		Niveau Image	
	Sens.	Spec.	Sens.	Spec.
0	33,48%	93,1%	66,7%	82,8%
1	53,8%	86,0%	93,2%	65,0%
2	56,8%	84,0%	93,3%	63,6%
3	57,1%	83,1%	93,7%	62,0%

Pour $depth = 0$, nous remarquons que la nouvelle agrégation des ROI permet d'augmenter la spécificité de 73% à 83% au niveau Image.

Nous remarquons aussi que la sensibilité augmente de 30% jusqu'à 57% au niveau ROI et de 53% jusqu'à 93,7% au niveau image en utilisant une profondeur de 3. Notre nouvelle méthode permet d'avoir un gain jusqu'à 40% de performance par rapport à la méthode initiale.

Si nous analysons la capacité de cette nouvelle méthode à augmenter la partie du polype présente dans une ROI nous obtenons les résultats suivants :

- 15,2% des ROI contiennent entre 100-90% d'un polype, soit un incrément de 3,7% ;
- 40,4% des ROI contiennent entre 90-50% d'un polype, soit un incrément de 8,8% ;
- 44,4% des ROI contiennent entre 50-1% d'un polype, soit une diminution de 12,5%.

La performance de trouver un polype au moins dans une ROI augmente avec la profondeur du suivi utilisée :

- pour $Depth = 0$, 60,7% des polypes ont été extraits dans une ROI qui contient au moins 50% du polype ;
- pour $Depth = 1$, 70,0% des polypes ont été extraits dans une ROI qui contient au moins 50% du polype ;

- pour $Depth = 2$, 71,1% des polypes ont été extraits dans une ROI qui contient au moins 50% du polype ;
- pour $Depth = 3$, 72,3% des polypes ont été extraits dans une ROI qui contient au moins 50% du polype.

Nous remarquons qu'augmenter la valeur de $Depth$ augmente les performances mais il faut stocker les $Depth$ images précédentes, ceci à un coût de stockage de $Depth * 1920 * 1080 * TaillePixel$ bits de mémoire. Si nous analysons les résultats, le gain entre $Depth = 1$ et $Depth = 3$ n'est que de 3,9%. Nous avons décidé de travailler avec $Depth = 1$ pour satisfaire la contrainte d'embarquabilité liée à l'intégration de ces traitements dans une CE.

4.3 Discussion

Dans ce chapitre nous avons évalué la méthode d'extraction automatique des ROI que nous utilisons et déterminé son impact dans la chaîne globale de l'acquisition de l'image jusqu'à la classification en utilisant une forêt floue. Au vue des résultats obtenus, nous avons modifié cette méthode afin d'améliorer les performances de détection. Pour cela nous avons introduit une nouvelle approche pour la détection automatique des polypes utilisant le suivi de mouvement.

Les résultats obtenus avec cette nouvelle méthode ont montré qu'en prenant en compte l'information spatiotemporelle d'une ROI, il est possible d'améliorer la détection de 53% jusqu'à 93,7% de sensibilité au niveau image.

Ce gain de sensibilité a un coût à payer au niveau de la spécificité au niveau d'image qui décroît pour atteindre un taux de 62%. Cette performance n'en reste pas moins intéressante par rapport à des méthodes de détection utilisée pour le diagnostic précoce du CCR, entre autres les méthodes de dépistage comme le test FIT.

Une des pistes pour améliorer la sensibilité et la spécificité est d'apprendre le classifieur pour reconnaître des ROI contenant plus ou moins de parties de polypes. Un ré-entraînement sera alors nécessaire, une des méthodes pourrait être de ré-entraîner le classifieur en utilisant les ROI extraites par la chaîne de traitement au lieu d'une l'extraction manuelle des ROI.

Nous avons testé cette dernière proposition et nous décrivons notre étude dans le chapitre suivant.

Chapitre 5

Entrainement du classifieur en utilisant les ROI extraites par la chaîne de traitement

I don't care that they stole my idea ...

I care that they don't have any of their own.

Nikola Tesla

Dans ce chapitre nous allons analyser la capacité de reconnaissance des polypes par une forêt floue pour la reconnaissance entraînée avec des ROI extraites automatiquement par notre chaîne de traitement d'images de coloscopies.

Pour réaliser cette analyse nous avons divisé les ROI extraites par la chaîne de traitement en cinq groupes :

- Des ROI contenant au moins 90% d'un polype.
- Des ROI contenant au moins 50% d'un polype.
- Des ROI contenant au moins 10% d'un polype.
- Des ROI qui sont soit trop grandes ou trop petites par rapport au polype qu'elles contiennent. Dans ce cas nous avons pris comme mesure quantitative du *trop* un facteur trois. Les ROI sont ici soit au moins trois fois plus grandes, soit au moins trois fois plus petites que le polype qu'elles contiennent.
- Des ROI ne contenant aucun polype.

À partir de la base d'apprentissage des images de coloscopie décrite dans le tableau 4.1, nous avons entraîné 3 forêts floues :

- une forêt entraînée à partir de ROI contenant au moins 90% d'un polype.

- une forêt entraînée à partir de ROI contenant au moins 50% d'un polype.
- une forêt entraînée à partir de ROI contenant au moins 10% d'un polype.

Chaque forêt est constituée de 10 arbres flous entraînés en utilisant le méthode de construction décrite dans le chapitre 3.3.

Chaque arbre a été entraîné à partir d'une base équilibrée constituée de 30 000 ROI, soit 15 000 ROI pour chaque classe.

Pour construire les bases, nous avons extrait les régions à l'aide de la vérité terrain et de notre chaîne de traitement de la manière suivante :

- À partir des ROI extraites pour notre chaîne de traitement et du masque binaire qui indique la localisation du polype, nous avons associé la classe "1" aux ROI contenant au moins 10% d'un polype.
- À partir des ROI extraites pour notre chaîne de traitement, seulement les ROI ne contenant pas de polype sont considérés comme classe "0".

Cette construction est illustrée dans la figure 5.1.

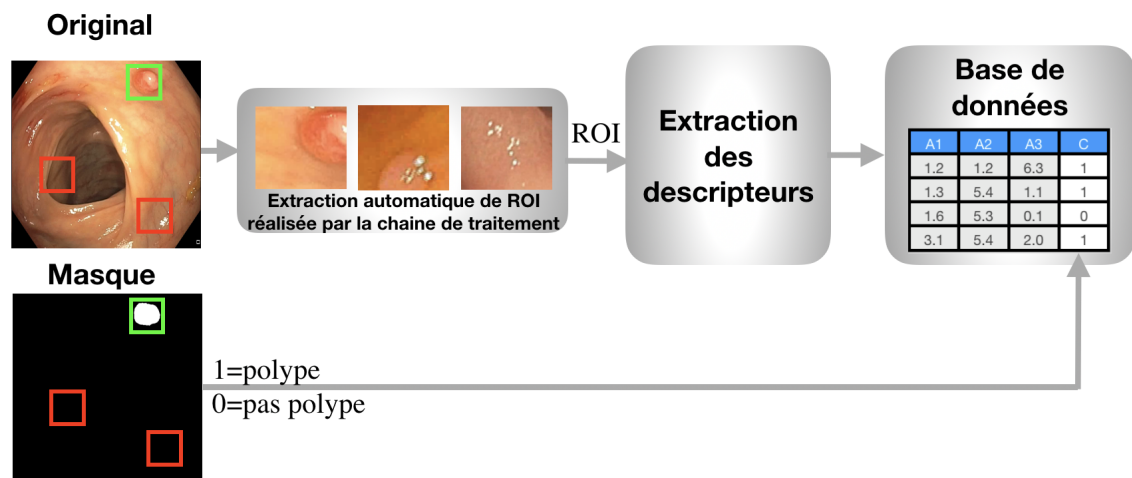


FIGURE 5.1 : Méthode proposée pour la construction de la base de données des ROI.

5.1 Expérimentations

Pour pouvoir observer les effets d'entraîner des arbres flous avec des ROI extraites automatiquement et ne contenant pas forcément l'intégralité d'un polype, nous n'avons pas dans un premier temps utilisé l'étape de suivi des ROI.

Les mesures de performances ont été réalisées sur les images de la base de test décrite dans le tableau 4.1.

Comme précédemment pour les performances au **niveau des ROI**, c'est le classifieur qui définit la classe, soit classe "1" pour présence de polype ou classe "0" pour l'absence de polype.

Pour les performances au **niveau image** nous avons utilisé la règle d'agrégation que nous avons déterminée au chapitre 4. Une image est classée classe "1" si elle contient au moins une ROI de classe "1" qui a été aussi classée "1" dans les N dernières images consécutives, avec $N = 3$ dans nos expériences.

Les métriques que nous avons utilisé pour cette évaluation sont comme précédemment la *sensibilité* = $VP/(VP + FN)$ et la *spécificité* = $VN/(VN + FP)$ avec comme définition pour prendre en compte la partialité de présence d'un polype dans une ROI :

- **VP**, c'est un ROI ou une image contenant au moins $P\%$ d'un polype classée comme "1" ;
- **FP**, c'est une ROI ou une image sans polype classée comme "1" ;
- **VN**, c'est une ROI ou une image sans polype classée comme "0" ;
- **FN**, c'est une ROI ou une image contenant au moins $P\%$ d'un polype classée comme "0".

Les résultats obtenus sont visibles sur la figure 5.2.

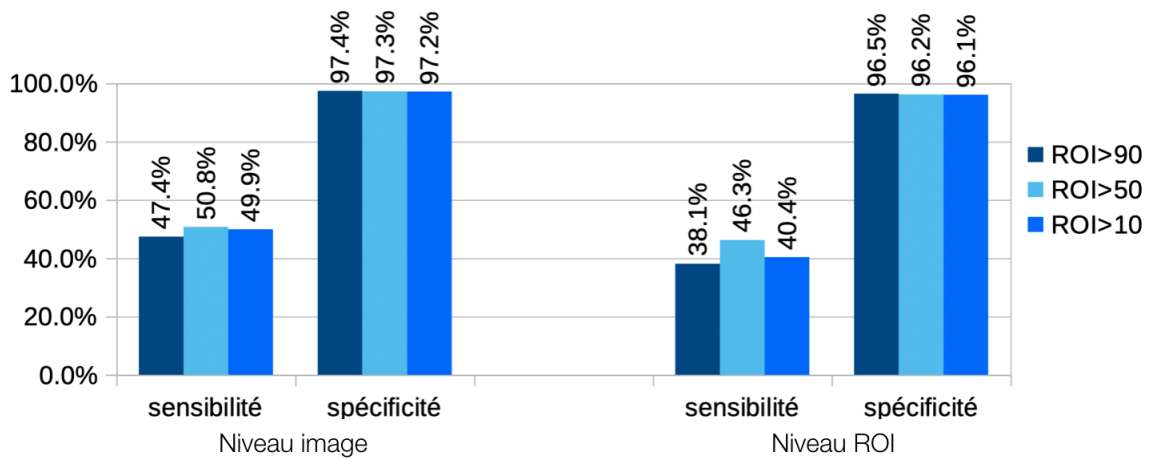


FIGURE 5.2 : Résultats de détection des trois forêts floues sur la base de test.

Ces résultats montrent des performances faibles pour les forêts qui ont été entraînées à partir des ROI contenant une partie d'un polype.

Pour comprendre ces résultats, nous avons évalué les trois forêts sur une base équilibrée constituée de 13 000 ROI, soit 6 500 ROI pour chaque classe, extraites de la base d'apprentissage décrite dans le tableau 4.1 pour vérifier si les forêts floues

ont appris à reconnaître des ROI contenant au moins un polype dans la base d'apprentissage.

Les résultats sont visibles sur la figure 5.3.

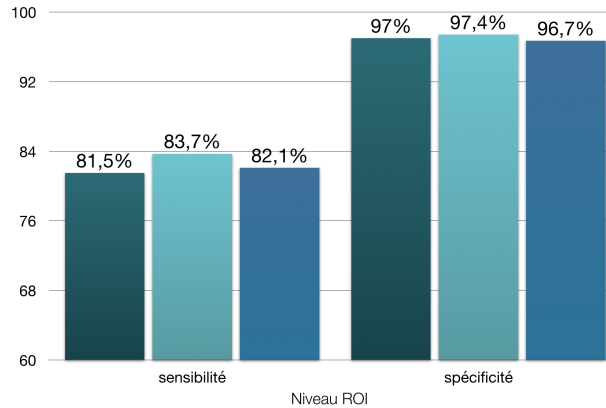


FIGURE 5.3 : Résultats de reconnaissance des trois forêts floues sur la base d'apprentissage.

Ces résultats indiquent que les forêts floues peuvent reconnaître des ROI contenant une partie d'un polype dès lors que nous utilisons des images de la base d'apprentissage, il est clair dans ce cas que le problème vient de la difficulté de ces forêts à généraliser et à reconnaître de nouveaux exemples.

La **généralisation** du modèle représente sa capacité, une fois entraîné, à effectuer des prédictions sur des données qu'il n'a jamais vues. C'est pourquoi les tests sont réalisés sur une autre base que la base d'apprentissage, la base de test, qui a pour vocation l'évaluation de cette capacité.

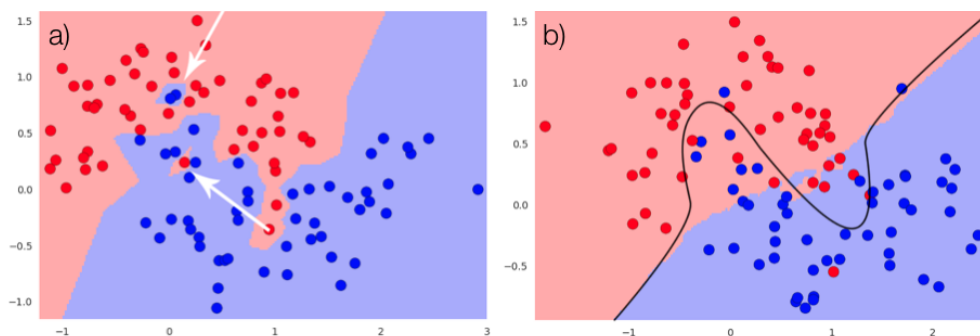


FIGURE 5.4 : Exemple a) d'un modèle qui a sur-appri, b) d'un modèle qui a sous-appri et un modèle de séparation quasi-idéal (ligne noire).

Il y peut y avoir deux raisons pour lesquelles notre modèle ne sais pas généraliser : soit il a sur-appri la base d'apprentissage soit il a sous-appri cette base. Ces phénomènes sont illustrés par la figure 5.4.

- Le **sur-apprentissage** désigne le fait que le modèle choisi est trop spécifique des données d'entraînement. Dans ce cas le modèle présente des performances très élevées pour la base d'apprentissage et des performances basses pour la base de test.
- Le **sous-apprentissage** désigne le fait que le modèle n'est pas assez complexe pour capturer le phénomène dans son intégralité. Cet effet est remarquable quand le modèle présente des performances moyennes dans la base d'apprentissage et des performances basses dans la base de test.

Au vue des résultats de la Figure 5.2 et de la Figure 5.3, nous sommes dans un problème de sous-apprentissage. Une solution pour résoudre ce problème est l'augmentation de données représentatives dans la base d'apprentissage. Une donnée représentative est une donnée que le classifieur ne classe pas correctement. En augmentant le nombre des données représentatives dans la base d'apprentissage le classifieur apprendra à classer correctement ces données représentatives. Plus spécifiquement, dans notre cas, il est nécessaire d'augmenter le nombre de **ROI** représentatives dans notre base d'apprentissage.

Pour augmenter la performance de détection de nos modèles en augmentant le nombre des **ROI** représentatives lors de l'entraînement des forêts floues, nous proposons une approche de Boosting décrite dans la section suivante.

5.2 Un approche de Boosting pour augmenter la performance de détection

Notre approche de Boosting est la suivante :

1. Une **base d'apprentissage et de test des ROI** est construite à partir des **ROI** extraites automatiquement de la **base d'apprentissage des images** (voir tableau 4.1) par notre chaine de traitement.
2. Le classifieur est entraînée à l'aide de la **base d'apprentissage des ROI**.
3. La chaine de traitement intégrant le classifieur est testée sur la **base d'apprentissage des images**. Les **ROI** qui n'ont pas été correctement classées par le classifieur correspondent aux **ROI** représentatives, puisque le classifieur n'a pas appris à reconnaître ces **ROI**. Ces **ROI** représentatives sont utilisées pour alimenter la **base d'apprentissage des ROI**. Le processus est réitéré à partir de l'étape 2 jusqu'à l'obtention de performances attendues ou d'un nombre d'itérations fixé.

Nous avons utilisé cette approche pour analyser la capacité de notre classifieur boosté à reconnaître des **ROI** contenant une partie d'un polype.

Avec cette approche nous cherchons à augmenter la performance de nos forêts floues en augmentant le nombre des **ROI** représentatives dans la **base d'apprentissage des ROI**.

5.2.1 Expérimentations

Pour chaque forêt, nous avons réalisé jusqu'à 10 itérations du processus que nous avons précédemment décrit.

Dans cette analyse, pour chaque itération, le classifieur est testé sur la **base de test des ROI** avec l'objectif d'analyser la capacité de notre classifieur à reconnaître les **ROI** et d'observer les progrès en performance. Et le classifieur est aussi testé sur la **base d'apprentissage des ROI** avec l'objectif d'analyser leur capacité de généralisation.

Aussi pour chaque itération, la chaîne de traitement intégrant la forêt floue est testée sur la **base de test des images** (voir tableau 4.1) avec l'objectif d'analyser l'impact du ré-entraînement de notre classifieur sur la performance au niveau des images.

Les résultats des performances obtenues au niveau **ROI** et au niveau image sont visibles sur les figures 5.5.a, 5.5.b et 5.5.c.

Nous remarquons que pour tous les cas, les performances de reconnaissance sur la **base d'apprentissage des ROI** baissent et sur la **base de test des ROI** augmentent. Ceci signifie que la méthode de Boosting a permis à notre modèle d'augmenter sa capacité de généralisation.

Nous constatons que la forêt entraînée pour reconnaître des **ROI** contenant au moins 50% d'un polype obtient des performances en sensibilité de 78,2% et en spécificité de 97,0% sur la base de test à la troisième itération

L'analyse globale des résultats au niveau image de la figure 5.5 nous montre que :

- pour la forêt entraînée avec des **ROI** contenant au moins 90% d'un polype (voir figure 5.5.a)), pour chaque itération, la sensibilité descend et la spécificité augmente ;
- pour la forêt entraînée avec des **ROI** contenant au moins 50% d'un polype (voir figure 5.5.b)), pour chaque itération, la sensibilité augmente et la spécificité augmente ;
- pour la forêt entraînée avec des **ROI** contenant au moins 10% d'un polype (voir figure 5.5.c)), pour chaque itération, la sensibilité augmente et la spécificité augmente ;

5.2. UN APPROCHE DE BOOSTING POUR AUGMENTER LA PERFORMANCE DE DÉTECTION

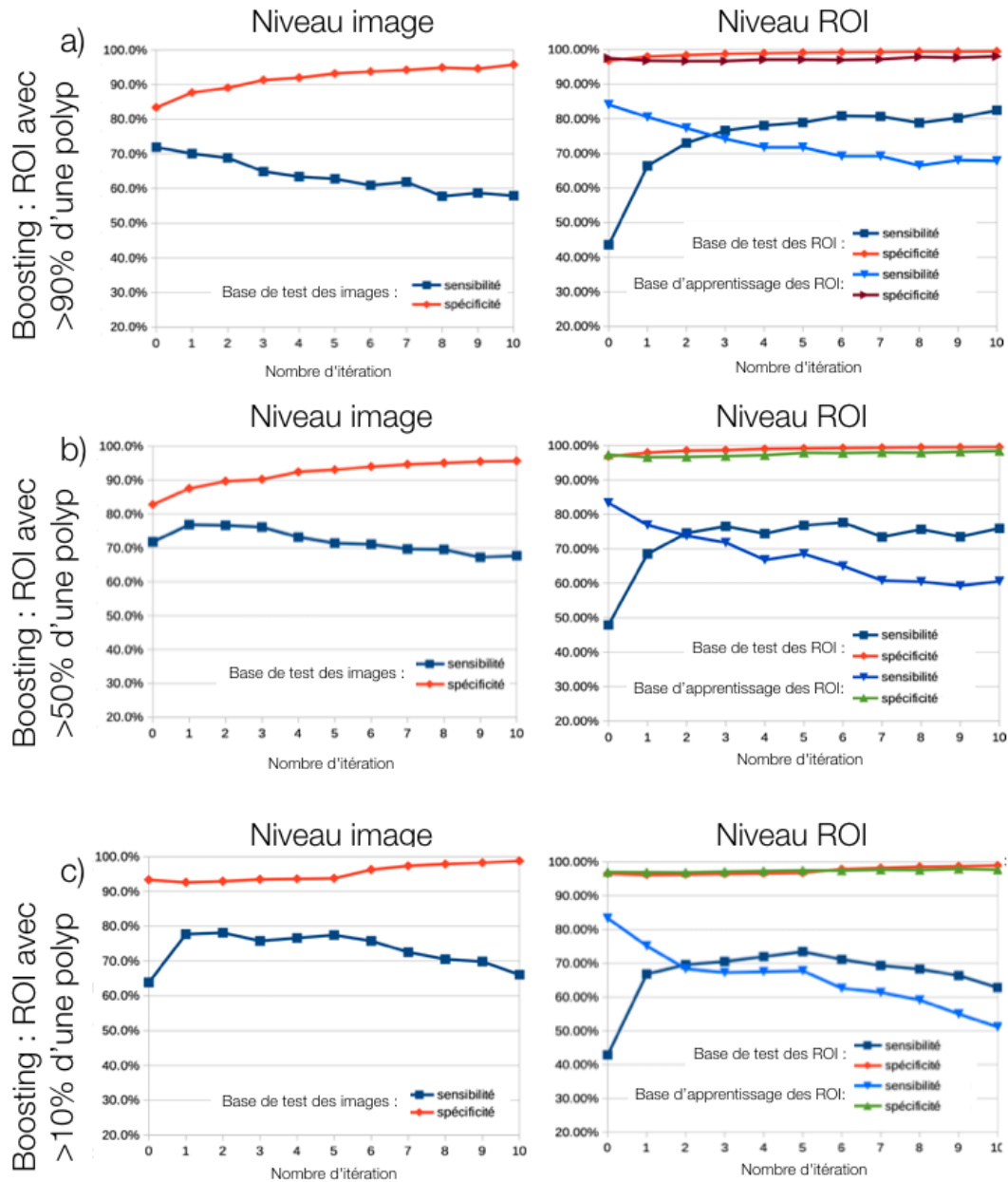


FIGURE 5.5 : Résultats des performances de reconnaissance des forêts floues boostées sur la base d'apprentissage et sur la base de test.

Pour comprendre cet effet, nous avons réalisé une analyse détaillée de la performance intrinsèque de trouver un polype au moins dans une [ROI](#) d'une image, nous a permis de constater que :

- pour 36,9% des images contenant un polype, le polype a été extrait dans des [ROI](#) qui contiennent au moins 90% de ce polype ;
- pour 69,4% des images contenant un polype, le polype a été extrait dans des [ROI](#) qui contiennent au moins 50% de ce polype ;
- pour 82,9% des images contenant un polype, le polype a été extrait dans des [ROI](#) qui contiennent au moins 10% de ce polype ;
- pour 6,7% des images contenant un polype, le polype a été extrait dans des [ROI](#) qui sont soit trop grandes ou soit trop petites par rapport au polype ;
- pour 10,4% des images contenant un polype, les polypes n'ont pas été insérés dans une [ROI](#).

À partir de l'analyse de ces résultats au niveau image de la figure 5.5, nous constatons que :

- pour la forêt entraînée avec des [ROI](#) contenant au moins 90% d'un polype la sensibilité descend puisque seulement 36,9% des polypes ont été extraits dans des [ROI](#) qui contiennent au moins de 90% d'un polype. Dans l'itération 0 de la figure 5.5.a, on peut voir que la forêt détecte 72% des polypes, cela signifie que la forêt a détecté $72\% - 37\% = 35\%$ des polypes extraits dans des [ROI](#) qui contiennent au moins de 90% de ce polype. Dans les itérations suivantes la forêt se spécialise pour ne détecter que les polypes extraits dans des [ROI](#) contenant au moins 90% et pas le reste. C'est pour cette raison que la sensibilité descend.
- pour la forêt entraînée avec des [ROI](#) contenant au moins 50% d'un polype la sensibilité et la spécificité augmente à chaque itération parce que la plupart des polypes, 70%, ont été extraits dans des [ROI](#) contenant au moins de 50% d'un polype.
- pour la forêt entraînée avec des [ROI](#) contenant au moins 10% d'un polype la performance augmente jusqu'à l'itération 2 de la figure 5.5.c et après descend. La forêt n'arrive pas à détecter les 82,9% polypes extraits dans des [ROI](#) contenant au moins de 10% d'un polype. Ceci montre qu'il faut une partie importante du polype contenue dans une [ROI](#) pour qu'il puisse être détecté.

Avec l'éclairage de ces résultats nous avons décidé d'adopter la deuxième forêt entraînée avec des [ROI](#) contenant au moins 50% d'un polype.

5.3 Mesures au niveau polype et patient

Au niveau médical, les performances des tests de dépistages et diagnostic sont mesurées en terme de sensibilité et spécificité au niveau **polype** et au niveau **patient**. Les résultats que nous avons montré jusqu'alors ont été mesurés au niveau **ROI** et au niveau **image**. Nous proposons de mesurer aussi les performances de la chaîne de traitement que nous proposons au niveau **polype** et autre au niveau **patient**.

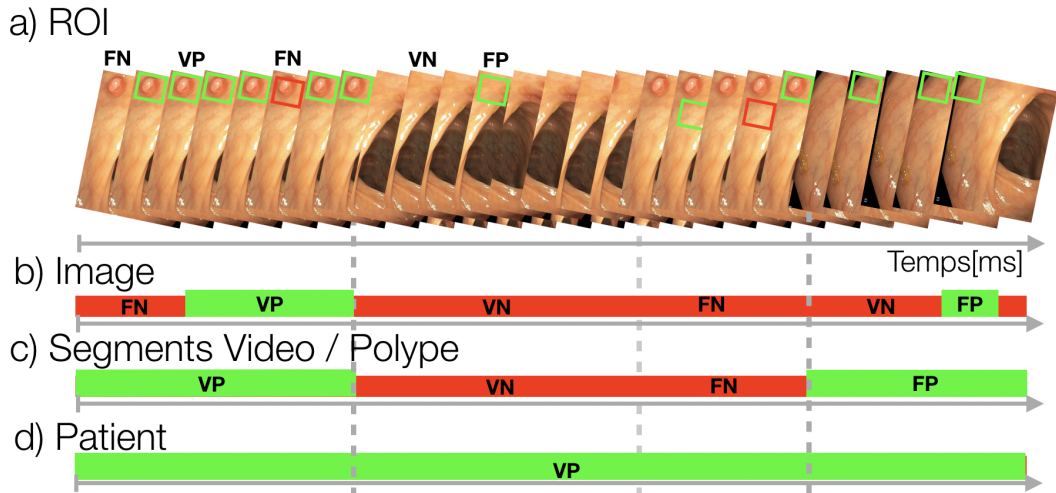


FIGURE 5.6 : Mesures au niveau polype et au niveau patient.

Ces métriques sont définies comme suit :

- Au niveau **polype**

- si dans un segment vidéo où il y a un polype, si celui-ci est détecté au moins une fois par le système cela correspond à un **VP**, s'il n'est pas détecté par le système cela correspond à un **FN** (voir figure 5.6.c).
- si dans un segment vidéo où il n'y a pas de polype, si le système détecte la présence d'un polype, cela correspond à un **FP**, si le système ne détecte pas la présence d'un polype cela correspond à un **VN** (voir figure 5.6.c).

- Au niveau **patient**

- dans une vidéo complète où il y a au moins un polype, si un polype est détecté au moins une fois par le système cela correspond à un **VP**, si aucun polype n'est détecté par le système cela correspond à un **FN** (voir figure 5.6.d).
- dans une vidéo complète où il n'y a pas des polype, si le système détecte la présence d'au moins un polype cela correspond à un **FP**, si le système ne détecte pas la présence de polype, cela correspond à un **VN** (voir figure 5.6.d).

L'ensemble de ces métriques au niveau **ROI**, au niveau **image**, au niveau **polype** et au niveau **patient** permettent d'évaluer la capacité de notre système à détecter un polype à la fois d'un point de vue technologie et d'un point de vue médical.

Nous avons évalué notre chaîne de traitement sur la **base de test des images** ASU-Mayo (voir tableau 4.1) qui est constituée de 10 films contenant un polype et 10 films ne contenant pas de polype. Dans les films avec polype, le polype apparaît plusieurs fois au cours de la vidéo. Pour ce test, nous avons considéré que chaque film appartient à un patient différent.

Au niveau polype, nous avons obtenu une sensibilité de 52% et une spécificité de 95%. Au niveau patient, nous avons obtenu une sensibilité de 100% et une spécificité de 70%.

Ces résultats montrent que :

- notre chaîne de traitement permet de détecter un polype sur un patient dès lors qu'il existe. Ce qui signifie qu'elle est capable de détecter ce polype sur au moins une image parmi l'ensemble des images d'une vidéo coloscopie.
- sur les 7 vidéo-coloscopies, parmi les 10, où il n'y a pas de polype, notre chaîne de traitement n'a fait aucune fausse détection.

Ces résultats montrent que notre système obtient une bonne performance, à la fois technologique et médicale. Ceci permet d'entrevoir sa capacité à être utilisé médicalement pour le diagnostic précoce du [CCR](#). Ces résultats sont à confirmer avec l'utilisation d'une base de données plus grande avec plus de patients.

Synthèse et discussion

Dans cette partie dans un premier chapitre nous avons proposé les arbres flous et des descripteurs de texture pour construire un modèle de classification pour la reconnaissance automatique des polypes. Nous avons évalué les performances de ce modèle sur des régions d'images de coloscopies susceptibles de contenir un polype, les [ROI](#). Pour cela nous avons entraîné le classifieur avec des [ROI](#) extraites manuellement. Les expérimentations ont montré que pour notre cas d'étude les arbres flous avec une classification binaire ont une bonne stabilité de 1% et une haute capacité de reconnaissance avec une sensibilité de 88% et une spécificité de 86%, ceci montre que le modèle de classification est assez performant. Les arbres flous avec une méthodes de classification floue ont montré aussi des bonnes capacités de reconnaissance, par contre ils sont moins stables.

Ensuite, pour améliorer la performance de reconnaissance nous avons proposé d'agréger de n modèle d'arbres flous Φ_n dans une seule forêt floue $\Psi\{\Phi_1, \Phi_2, \dots, \Phi_n\}$. Les expériences ont confirmé l'amélioration des performances de classification, dans le cas d'une classification binaire la sensibilité augmente de 88% à 93% et la spécificité s'améliore de 86% à 91%. Les méthodes de classification floues arrivent à une sensibilité de 90% et une spécificité de 86%.

Ces expériences nous ont permis aussi de valider les descripteurs de texture et de luminosité comme des descripteurs discriminants dans la reconnaissance des polypes.

Nous avons ensuite présenté notre méthode d'extraction automatique des [ROI](#), le pré-traitement nécessaire pour leur classification. La transformée de Hough à été choisi comme filtre permettent d'extraire les contours circulaires. Les expériences qui ont été menées en utilisant ce pré-traitement avec les classifieurs précédemment décrits ont montré une forte diminution de la sensibilité de 93% à 30% au niveau des [ROI](#).

L'analyse de ces résultats a permis d'identifier un problème dans l'extraction automatique proposée où les polypes ne sont pas forcément centrés et entiers dans les [ROI](#). Les mesures ont montré aussi un lien clair entre les performances en sensibilité et la proportion du pourcentage du polype contenu dans une [ROI](#).

Nous avons réalisé une analyse de la performance au niveau image. Nous avons remarqué que la règle d'inférence utilisée pour agréger les [ROI](#) et donner une déci-

sion image n'était pas efficace. Nous avons proposé une nouvelle règle d'agrégation temporelle pour dépasser ces limites.

Afin d'améliorer les performances de la chaîne de traitement nous avons conçu une nouvelle méthode pour optimiser l'extraction des [ROI](#). Cette solution utilise un suivi temporel des [ROI](#) dans une séquence d'images en estimant le mouvement de ces régions. Les expériences ont montré que cette prise en compte de l'information spatio-temporelle d'une [ROI](#) améliore la détection de 53% jusqu'à 93,7% en sensibilité au niveau image, mais avec une spécificité de 62%.

Ensuite, afin d'améliorer les performances de classification indépendamment du pourcentage de polype présent dans une [ROI](#), nous avons ré-entraîné le classifieur en utilisant les [ROI](#) extraites par la chaîne de traitement en lieu et place de l'extraction manuelle des [ROI](#).

Nous avons utilisé une approche de Boosting pour analyser la capacité des forêts floues à reconnaître des [ROI](#) contenant une partie d'un polype. Les résultats ont montré une augmentation dans les performances de détection jusqu'à une sensibilité de 84,3% et une spécificité et de 88,3% au niveau image.

Finalement, nous avons proposé des nouvelles métriques pour évaluer la performance de notre méthode au niveau polype et au niveau patient pour permettre une interprétation médicale des performances. Les résultats ont montré que notre méthode obtient une très bonne performance sur la base de données ASU Mayo.

Pour notre cas d'étude, cette chaîne de traitement obtient une performance de détection largement supérieure à celle des tests utilisés actuellement pour le diagnostic précoce du [CCR](#). Il est nécessaire maintenant d'étudier son intégration dans une [CE](#) ce qui est l'objet de la prochaine partie.

Troisième partie

MISE EN ŒUVRE MATÉRIELLE

Introduction	87
Profilage de la chaine de traitement sur un processeur embarqué	88
6 État de l’art : transformée de Hough et ses implémentations	92
6.1 Algorithme original de la transformée de Hough	93
6.2 Algorithme de la transformée d’Hough Unidimensionnelle (1DHT) . .	93
6.3 Synthèse et analyse	96
7 Mise en œuvre de l’algorithme 1DHT en temps réel	98
7.1 Accélération de calcul des centres candidates	98
7.2 Architecture pour la mise en œuvre de l’algorithme 1DHT modifié . .	103
7.2.1 Calcul des coordonnées en x	104
7.2.2 Calcul du rayon	108
7.3 Mise en œuvre de l’architecture du 1DHT IP sur un FPGA	109
7.3.1 Modules Calcul des x-coordonnées et Calcul des y-coordonnées	109
7.3.2 Module Calcul du rayon	109
7.4 Mise en œuvre de l’architecture du 1DHT IP dans un Système sur Puce	112
8 Optimisation de l’IP 1DHT en surface	118
8.1 Optimisation du calcul des centres candidats	118
8.2 Architecture proposée pour la mise en œuvre du calcul des coordon- nées en x	121
8.3 Intégration de l’IP 1DHT-optimisée sur FPGA	125
8.4 Intégration de notre IP 1DHT-optimisée dans une système sur puce .	125
8.5 Estimation de la consommation de l’IP optimisée	127
8.5.1 Avec un FPGA Spartan 7 XC7S15	127
8.5.2 Avec un ASIC en technologie 65 nm ST	128
Synthèse et discussion	129

Introduction

Pour concevoir le SVSoC que nous désirons intégrer dans une CE, il est nécessaire de s'assurer de l'embarquabilité de la chaîne de traitement définies dans la partie II. Pour cela il faut prendre en compte de toutes les contraintes non fonctionnelles liées à cette intégration. Parmi ces contraintes il y a la latence du calcul qui doit être inférieure à 40 ms, ce qui correspond à une cadence standard vidéo de 25 images par seconde, la surface d'intégration de $8 * 8mm^2$ puisque le diamètre d'une CE est d'environ 1 cm et le budget énergétique. Pour ce dernier si nous considérons deux batteries de type pile bouton standards Energizer EPX76 (voir Tableau 1.5) de caractéristiques 1,55 V pour 188 mAh nous disposons d'un budget de 582 mWh.

Nous allons tout d'abord réaliser le profilage de notre chaîne de traitement avec une mise en œuvre sur un processeur embarqué afin de mesurer et d'analyser ses performances. Ce profilage va nous permettre aussi d'identifier les parties de la chaîne qui nécessitent une accélération matérielle.

Nous faisons ensuite un état de l'art portant sur la ou les partie(s) identifié(es) comme critiques en terme de latence et nous analysons les différentes mise en œuvre existantes.

Nous proposons à la suite de cet état de l'art une architecture matérielle pour l'accélération des parties critiques avec une mise en œuvre et une validation sur un système sur puce.

Pour finir, nous optimisons cette architecture en considérant la surface disponible dans une CE.

Profilage de la chaine de traitement sur un processeur embarqué

Pour réaliser nos analyses, nous avons choisi de mettre en œuvre notre chaîne en utilisant un système sur puce, le Zynq-7000 XC7Z020 intégré sur une carte de développement de la société Digilent, la carte ZedBoard.

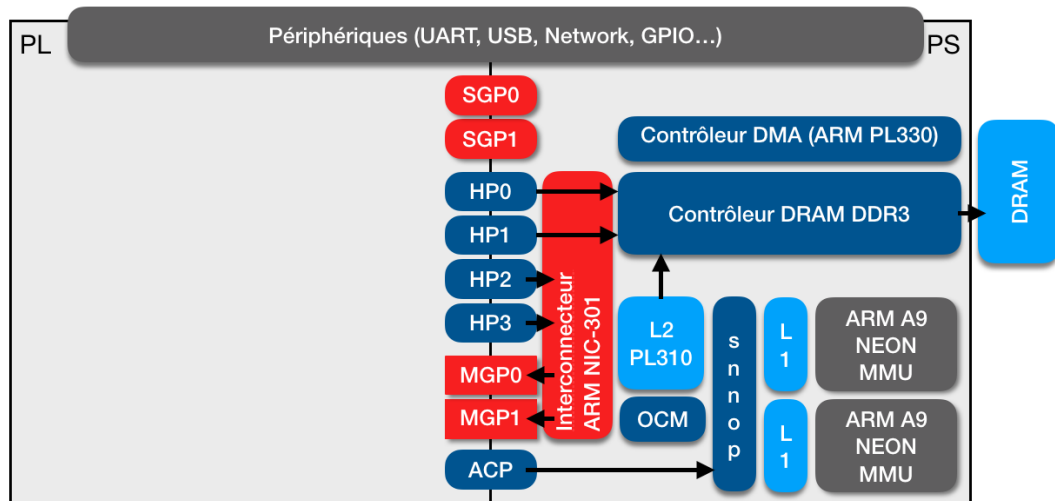


FIGURE 5.7 : Architecture Zynq-7000 [9].

Le Zynq, visible sur la figure 5.7, est un circuit proposé par le constructeur Xilinx qui est constitué de deux parties :

- une partie de logique programmable pour la mise en œuvre d'architectures matérielles. Cette partie comprend 53,200 LUT, 4,9Mb en BRAM et 220 Blocs DSP ;
- une partie système avec un processeur, où il est possible d'exécuter un programme logiciel [13].

Ces deux parties communiquent à l'aide d'un bus d'interconnexion de type AMBA AXI4[11].

La partie de logique programmable est constituée d'une architecture de type artix7 et la partie système est constituée d'un processeur double-cœur [Advanced RISC Machine \(ARM\)](#) Cortex A9, d'un contrôleur de mémoire DRAM, de mémoires caches de niveau L1 et L2, d'un contrôleur de cache SNNOP et de différents périphériques tels qu'une UART, un port USB, un port Ethernet, des GPIO.

La carte de développement ZedBoard, visible sur la figure 5.8, est une carte proposée par la société Digilent [9] qui intègre un circuit Zynq et tous les périphériques qui lui sont interconnectés. Elle permet un développement logiciel et matériel d'applications intégrées au sein du circuit Zynq.

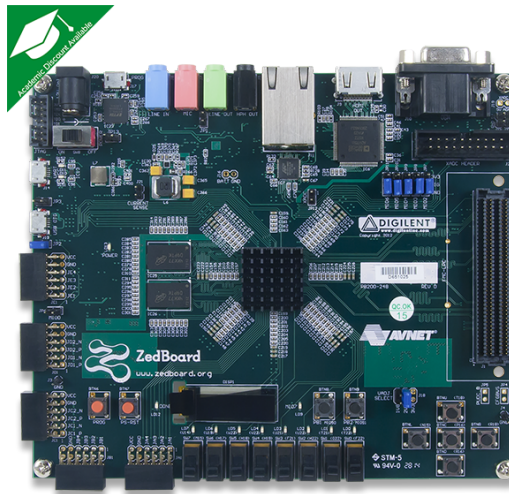


FIGURE 5.8 : Carte de développement Zedboard (source : www.digilentinc.com).

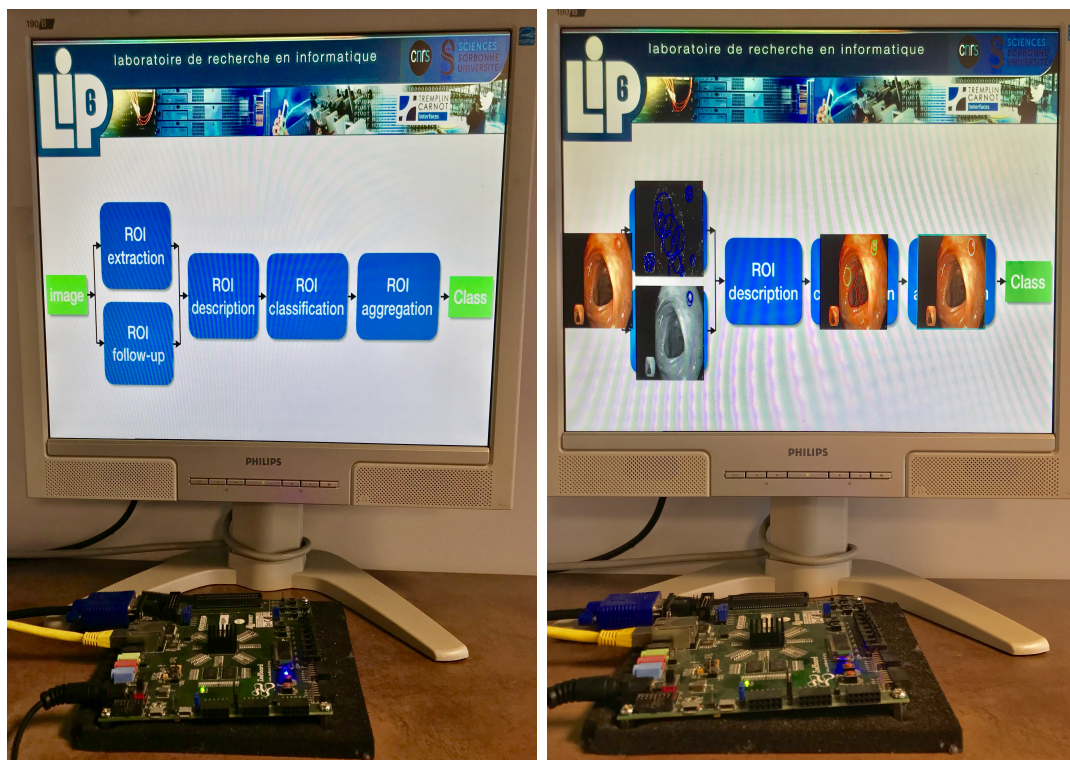


FIGURE 5.9 : Intégration dans la carte de développement Zedboard.

Nous avons choisi d'intégrer notre chaîne au sein du Zynq, pour faire un prototype rapide des architectures numériques que nous avons développé. L'architecture Zynq possède une interface flexible entre le processeur et la partie logique programmable. Cela nous permet facilement de réaliser un système possédant des parties logicielles intégrées sur le processeur ARM Cortex A9 et des parties matérielles réalisées sous forme d'accélérateurs matériels et intégrés dans la logique programmable. La méthodologie suivie ici est de tout d'abord faire une mise en œuvre purement logicielle et de migrer en matériel les parties identifiées comme chronophages.

Nous avons conçu un système d'expérimentation à l'aide de la carte de développement Zedboard qui est visible sur la figure 5.9.

Pour réaliser la mise en œuvre logicielle nous avons utilisé la bibliothèque optimisée OpenCV. Notre chaîne de traitement a été exécutée sur le processeur ARM Cortex A9 cadencé à la fréquence de 667 MHz.

Pour mesurer le temps d'exécution nous avons utilisé la librairie *xtime.h* et le *TIMER0* intégré dans le processeur ARM Cortex A9.

Nous avons exécuté notre chaîne de traitement sur 1234 images de coloscopie en FHD (1920x1080 pixels) et en couleurs codées en RGB afin de mesurer le temps d'exécution maximum de chaque étape. Les mesures réalisées sont visibles dans le tableau 5.1.

Les images proviennent des vidéos 11 et 12 de la base de données ASU-Mayo [89] décrite dans le tableau 2.1 de la section 2. Ces vidéos ont été sélectionnées aléatoirement dans la base de données.

Nous avons limité le nombre maximum des ROI à 13, ce qui correspond au nombre de ROI moyen qui a été mesuré dans le chapitre 4.

TABLE 5.1 : Profil des temps d'exécution de la chaîne de traitement pour des images FHD.

Partie de la chaîne	Calcul	Latence
Extraction des ROI	Niveau de Gris	35,9ms
	Filtre Médian	291,8ms
	Filtre Canny	233,5ms
	Transformée de Hough	764,9ms
Suivi des ROI	Correspondance par bloc	68,7ms
Description des ROI	Matrice de co-occurrence	15,9ms par ROI
	Descripteurs de Haralick	91,5ms par ROI
Classification des ROI	Arbre flou	1,9ms x N arbres par ROI
Agrégation des ROI	Agrégation	0,09ms

Ce profil d'exécution nous indique que notre chaîne a un temps d'exécution total de 3,0229 secondes pour 13 ROI dans une image FHD. Ce temps est 77 fois plus grand que notre objectif qui est de traiter une image toutes les 40 ms.

Nous pouvons constater aussi en regardant les données du tableau 5.1, qu'il y a deux parties très chronophages :

- la transformée de Hough dans l'étape d'extraction des [ROI](#);
- les descripteurs de Haralick dans l'étape de description des [ROI](#).

Ces résultats excluent une optimisation logicielle pour une mise en œuvre embarquée, en utilisant par exemple un jeu d'instruction SIMD comme le jeu d'instruction NEON du processeur [ARM](#) Cortex A9, et nous oblige à concevoir une architecture matérielle pour optimiser la latence de calcul de ces deux parties. Cette architecture sera intégrée dans un [FPGA](#) ou dans un [ASIC](#).

Nous avons décidé de commencer pour le point plus critique, le calcul de la transformée d'Hough dont la latence est 20 fois plus grande que notre objectif de 40 ms.

Dans le prochain chapitre, nous allons réaliser un état de l'art de la mise en œuvre matérielle de la transformée de Hough en analysant les performances en terme de temps d'exécution et de surface utilisée afin de les comparer à nos besoins pour l'intégration dans une [CE](#).

Chapitre 6

État de l'art : transformée de Hough et ses implémentations

*Imagination is more important than knowledge. Knowledge is limited.
Imagination encircles the world.*

Albert Einstein

Un état de l'art a été réalisé pour analyser la mise en œuvre matérielle de la transformée de Hough. Les travaux récents que nous avons intégré dans cet état de l'art considèrent tous une implantation finale dans un [FPGA](#).

Une étude des méthodes de calcul de la transformée de Hough est décrite dans [21] et dans [71].

Tous les travaux se focalisent essentiellement sur l'accélération du processus de vote de la transformée de Hough. L'objectif est d'accumuler des points d'intersection dans l'espace des paramètres de Hough qui correspondent au nombre de cercles possibles. Un vote est réalisé pour trouver le maximum local, en terme d'accumulation de points, qui sera considéré comme un cercle réel. Pour mettre en œuvre ce processus il existe au moins deux approches différentes :

- une première approche utilise l'algorithme original de la transformée de Hough où des équations paramétriques du cercle sont utilisées pour trouver le centre et le rayon des cercles [41], [38], [83], [52], [59] ;
- une seconde approche utilise une version modifiée de la transformée de Hough appelée transformée de Hough Unidimensionnelle ([One Dimensional Hough Transform \(1DHT\)](#))[100].

6.1 Algorithme original de la transformée de Hough

Pour la mise en œuvre de l'algorithme original de la transformée de Hough, Elhossini et al [38] ont démontré que l'usage d'une mémoire externe pour stocker les résultats des votes est limitée par la bande passante de transfert des données. Ils proposent une architecture de mémoire interne pour la mise en œuvre du processus de vote, avec cette approche la taille de la mémoire nécessaire est de 256Kbit et les calculs sont distribués dans 32 unités de calcul, chaque unité a accès à sa propre mémoire en utilisant deux buffers. Cette structure permet deux opérations de lecture/écriture en parallèle pour chaque unité de calcul.

Ferhat et al [41] proposent une mise en œuvre de la transformée de Hough pour la détection des cercles basée sur l'utilisation de l'algorithme CORDIC [19], le principal aspect de ce travail est qu'il détecte un seul cercle. La même approche est utilisée pour Kumar et al [58] où un accélérateur matériel est proposé pour la localisation d'iris.

Irwansyah et al [52] ont proposé un algorithme pour la mise en œuvre de la transformée de Hough combiné avec un algorithme de regroupement de graphes pour la détection de plusieurs cercles.

6.2 Algorithme de la transformée d'Hough Unidimensionnelle (1DHT)

Goneid et al [45] ont proposé, en 1997, une version modifiée de la transformée de Hough pour la détection de plusieurs cercles. Cette méthode peut extraire des cercles et des ellipses non superposés à partir d'une image de contour, même lorsque les contours ne sont pas réguliers. Elle est représentée par le pseudocode de l'algorithme 2 et peut être résumée en trois étapes (voir figure 6.1) :

1. À partir d'une image de contour, deux histogrammes sont créés :
 - Lignes [1-7] : un histogramme des points milieux des segments formés par tous les couples de points de contours sur une même ligne. Cet histogramme est appelé H_x ;
 - Lignes [8-14] : un histogramme des points milieux des segments formés par tous les couples de points de contours sur une même colonne. Cet histogramme est appelé H_y ;
2. Les maximums locaux combinés des histogrammes H_x et H_y deviennent la coordonnée en x et la coordonnée en y de possibles centres de cercles, ou d'ellipses, ces coordonnées sont dénotées (i, j) . Elle deviennent les coordonnées de

centres candidats à devenir un cercle, ou une ellipse. A partir de ces coordonnées :

- Lignes [17-22] : un histogramme des distances entre tous les points de contour et les centres candidats est calculé. Cet histogramme est appelé H_r . La valeur de la distance r la plus votée de cet histogramme, pour chaque centre candidat (i, j) , devient son rayon ;
- Lignes [22-26] : si la valeur $H_r(i, j, r)$ est plus grande que $f * 4\sqrt{2}r$ alors le centre candidat (i, j) et son rayon sont identifiés comme étant un cercle d'équation $\|(x, y) - (i, j)\| = r$.

Algorithm 2 Algorithme de la [1DHT](#)

```

1: for chaque ligne  $i$  de 1 à  $H$  do
2:   for chaque point de contour  $I(i, j)$  avec  $j$  de 1 à  $W - 1$  do
3:     for chaque point de contour  $I(i, k)$  avec  $k$  de  $j + 1$  à  $W$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + 1$ 
5:     end for
6:   end for
7: end for
8: for chaque colonne  $i$  de 1 to  $W$  do
9:   for chaque point de contour  $I(j, i)$  avec  $j$  de 1 à  $H - 1$  do
10:    for chaque point de contour  $I(k, i)$  avec  $k$  de  $j + 1$  à  $H$  do
11:       $H_y(\frac{j+k}{2}) = H_y(\frac{j+k}{2}) + 1$ 
12:    end for
13:  end for
14: end for
15: calcul des maximums locaux  $H_x(i)$  de  $H_x$ 
16: calcul des maximums locaux  $H_y(i)$  de  $H_y$ 
17: for chaque maximum local  $H_x(i)$  de  $H_x$  do
18:   for chaque maximum local  $H_y(j)$  de  $H_y$  do
19:     for all points de contour  $I(k, m)$  de l'image do
20:        $H_r(i, j, \|I(k, m) - (i, j)\|) = H_r(i, j, \|I(k, m) - (i, j)\|) + 1$ 
21:     end for
22:     if valeur maximale de  $H_r(i, j, r) > f * (4\sqrt{2}r)$  then
23:        $\|(x, y) - (i, j)\| = r$  est une cercle
24:     end if
25:   end for
26: end for

```

Cette méthode peut être mise en œuvre facilement puisque chaque paramètre est accumulé dans son propre espace de paramètres.

Zhou et al. [100] ont proposé une architecture qui met en œuvre cet algorithme au sein d'un FPGA. Ils ont choisi d'utiliser des entiers codés sur 9 bits, pour les deux histogrammes H_x et H_y et des entiers codés sur 17 bits pour l'historgramme H_r .

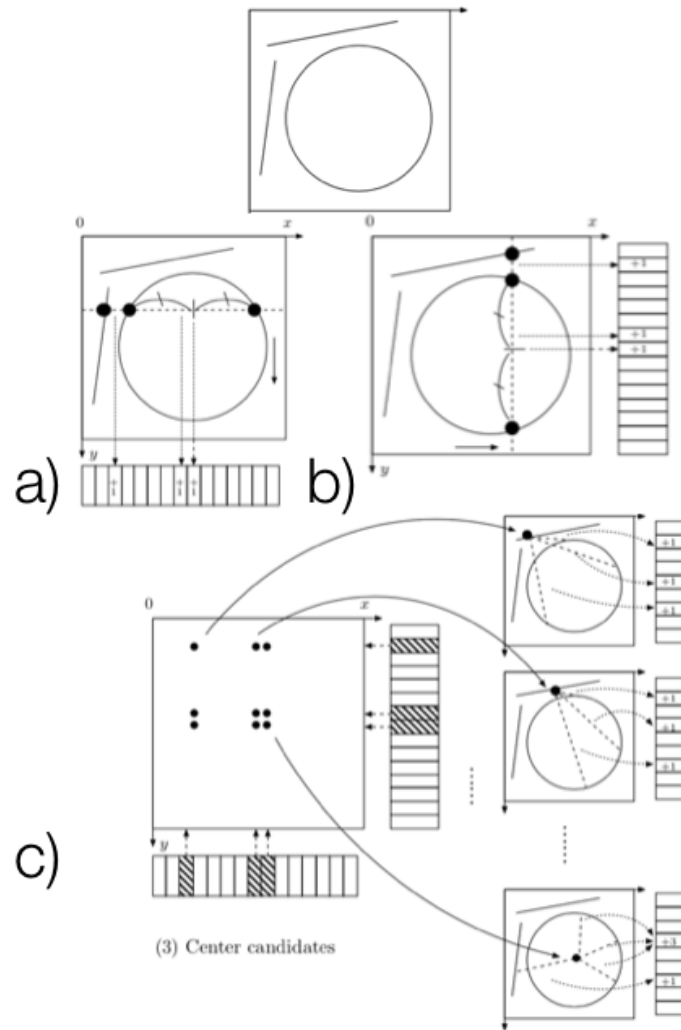


FIGURE 6.1 : a) Calcul des x-coordonnées. b) Calcul des y-coordonnées. c) Calcul des rayons.

Ils utilisent des mémoires internes du FPGA, des BRAM, pour mettre en œuvre ces histogrammes en utilisant 100 modules de vote en parallèle. Chaque unité de vote permet le calcul de points milieux entre deux points de contours. 10 cordonnées en x et 10 cordonnées en y sont calculées à partir des maximums locaux des histogrammes H_x et H_y . 100 centres candidats ont été créés à partir de la combinaison des cordonnées en x et des cordonnées en y. Chaque cordonnée en x et chaque cordonnée en y est codée avec des entiers sur 9 bits. Chaque rayon est codé avec des entiers sur 13 bits.

6.3 Synthèse et analyse

Les performances en terme de latence et les caractéristiques des différents travaux trouvés dans la littérature sont listées dans la table 6.1.

TABLE 6.1 : Performances de l'état de l'art pour la transformée de Hough.

Auteurs	FPGA	Ressources BRAM [Kb] /DSP	Résolution [pixel]	Latence [ms] /fréq. [MHz]	Débit [$\frac{Mpxl}{s}$]	Latence [ms] extrapolé à FHD	BRAM [Kb] extrapolé à FHD
Ferhat et al, 2012[42]	Virtex 2	144 / -	1 cercle	- / 66,9	-	-	-
Elhossini et al, 2012[38]	Virtex 4	256 / -	800x600	33,3 / -	14,40	144	1105
Zhou et al, 2014[100]	Virtex 7	5562/398	400x400	5,3/181,8	29,97	69,18	72084
Seo et al, 2015[83]	Kintex 7	- / -	640x480	10/250	30,72	67,5	-
Irwansyah et al, 2015[52]	Virtex 4	512 / -	512x512	7,81/135	33,55	61,8	4050
Kumar et al, 2017[58]	Artix 7	1872 / -	320x240	3,46/203	22,20	93,42	50544

À partir de cet état de l'art nous pouvons observer :

1. que toutes les mises en œuvre [41], [38], [100], [83], [52], [58] utilisent des blocs de mémoire interne ou Blocs RAM (BRAM) ;
2. que tous les travaux se focalisent essentiellement sur l'accélération du processus de vote de la transformée de Hough. Cette étape du calcul est chronophage et nécessite un espace mémoire important, pour savoir la quantité de mémoire nécessaire pour notre cas, nous avons extrapolé avec une simple mise à l'échelle les résultats pour des images de résolution [FHD](#) (1920x1080 pixels) ;
3. qu'il n'existe pas de mise en œuvre d'architecture pour des images de résolution [FHD](#). Afin de pouvoir savoir si les solutions existantes pourraient convenir pour ce type d'image, nous avons extrapolé avec une simple mise à l'échelle les résultats pour des images de résolution [FHD](#) (1920x1080 pixels). Il en ressort qu'aucun des travaux ne permet d'obtenir une latence inférieure à 40 ms.

Aucun des travaux de l'état de l'art ne permet d'obtenir un temps de calcul de 40ms pour des images [FHD](#), nous allons concentrer nos efforts sur la conception d'une nouvelle architecture pour accélérer le calcul de la transformée de Hough pour ce type d'image.

Avant de débiter cette conception, il est nécessaire de choisir entre les deux algorithmes décrits dans les sections 6.1 et 6.2.

Si nous nous référons aux travaux présentés dans [54, 67], il est indiqué que les polypes peuvent être décrits par une forme de protubérances et détectés à l'aide de la courbure de ses contours en cherchant à identifier sur l'image les formes circulaires ou elliptiques. Des travaux d'analyse des images de polypes à partir de coloscopies [78], ont permis aussi d'observer que les polypes n'ont pas toujours une forme circulaire ou elliptique régulière, cette irregularité dépend du niveau de bruit, de la qualité et de la résolution de l'image. Il apparaît dès lors que l'algorithme de Goneid est plus adapté puisqu'il peut extraire des cercles/ellipses non superposés à partir d'une image de contour, même lorsque les contours ne sont pas réguliers due à la présence de bruit aléatoire. En ajoutant à cela que l'algorithme de Goneid est simple à mettre en œuvre parce que chacun de ses paramètres est accumulé dans son propre espace de paramètres [100], nous avons choisi d'étudier cet algorithme pour réaliser une architecture efficace avec comme ambition de calculer la transformée de Hough en moins de 40 ms pour des images [FHD](#).

Chapitre 7

Mise en œuvre de l’algorithme 1DHT en temps réel

The biggest risk is not taking any risk... In a world that is changing really quickly, the only strategy that is guaranteed to fail is not taking risks

Mark Zuckerberg

Nous avons avant toute réalisation, étudié l’algorithme [1DHT](#) proposé par Goneid et al. [45] pour déterminer comment il peut être embarqué dans une [CE](#).

7.1 Accélération de calcul des centres candidates

Notre étude de la construction des histogrammes H_x et H_y , étapes [1-7] et [8-14] de l’algorithme 2, nous a permis de proposer une solution pour optimiser leur calcul.

Pour le calcul de l’histogrammes H_x , étapes [1-7] de l’algorithme 2, nous avons fait les optimisations suivantes :

- l’image peut être lue de manière continue en prenant en compte tous les points de l’image au lieu de ne prendre en compte que les points de contours, comme cela est décrit dans l’algorithme 3.
- Le vote total correspondant aux points milieux entre la colonne j et k à la fin de chaque boucle j peut être obtenu en changeant l’ordre des boucles i , j et k comme cela est décrit dans l’algorithme 4.
- Dans les boucles j et k de l’algorithme 4, la colonne j est multipliée par les $W - j - 1$ colonnes suivantes. Nous pouvons limiter le nombre de calcul en

Algorithm 3 Modification de l'algorithme 1DHT : optimisation de la lecture des points lus

```

1: for chaque ligne  $i$  de 1 à  $H$  do
2:   for chaque point  $I(i, j)$  avec  $j$  de 1 à  $W - 1$  do
3:     for chaque point  $I(i, k)$  avec  $k$  de  $j + 1$  à  $W$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + I(i, j) * I(i, k)$ 
5:     end for
6:   end for
7: end for

```

Algorithm 4 Modification de l'algorithme 1DHT : changement de l'ordre des boucles pour le point milieu

```

1: for chaque colonne  $j$  de 1 à  $W - 1$  do
2:   for chaque colonne  $k$  de  $j + 1$  à  $W$  do
3:     for chaque ligne  $i$  de 1 à  $H$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + I(i, j) * I(i, k)$ 
5:     end for
6:   end for
7: end for

```

fixant une recherche de cercle d'un diamètre maximal D_m . Il est alors suffisant de multiplier la colonne j avec les D_m colonnes suivantes, comme cela est décrit dans l'algorithme 5.

Algorithm 5 Modification de l'algorithme 1DHT : limitation du nombre de multiplications grâce au paramètre D_m

```

1: for chaque colonne  $j$  de 1 à  $W - D_m$  do
2:   for chaque colonne  $k$  de  $j + 1$  à  $j + D_m$  do
3:     for chaque ligne  $i$  de 1 à  $H$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + I(i, j) * I(i, k)$ 
5:     end for
6:   end for
7: end for

```

En appliquant les mêmes optimisations pour le calcul de l'histogrammes H_y , étapes [8-14] de l'algorithme 2, nous obtenons un algorithme équivalent, présenté dans l'algorithme 6, dans lequel nous obtenons les mêmes histogrammes H_x et H_y que l'algorithme original 2.

L'algorithme initial 2 a besoin, pour la construction des histogrammes H_x et H_y , d'une lecture entière de l'image, cela implique que les maximums locaux des histogrammes H_x et H_y peuvent être calculés une fois seulement après avoir lu toute l'image. Il est donc nécessaire de mémoriser les deux histogrammes H_x et H_y .

L'avantage de notre nouvel algorithme 6, dans la construction des histogrammes H_x et H_y , est que les valeurs d'accumulation de l'histogramme H_x sont calculées pour

Algorithm 6 Algorithme [1DHT](#) modifié proposé

```

1: for each column  $j$  from 1 to  $W - Dm$  do
2:   for each column  $k$  from  $j + 1$  to  $j + Dm$  do
3:     for each row  $i$  from 1 to  $H$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + I(i, j) * I(i, k)$ 
5:     end for
6:   end for
7: end for
8: for each row  $j$  from 1 to  $H - Dm$  do
9:   for each row  $k$  from  $j + 1$  to  $j + Dm$  do
10:    for each col  $i$  from 1 to  $W$  do
11:       $H_y(\frac{j+k}{2}) = H_y(\frac{j+k}{2}) + I(j, i) * I(k, i)$ 
12:    end for
13:  end for
14: end for
15: calcul des maximums locaux  $H_x(i)$  de  $H_x$ 
16: calcul des maximums locaux  $H_y(i)$  de  $H_y$ 
17: for chaque maximum local  $H_x(i)$  in  $H_x$  do
18:   for chaque maximum local  $H_y(j)$  in  $H_y$  do
19:     for all edge  $I(k, m)$  in image do
20:        $H_r(i, j, ||I(k, m) - (i, j)||) = H_r(i, j, ||I(k, m) - (i, j)||) + 1$ 
21:     end for
22:     if max value of  $H_r(i, j, r) > f * (4\sqrt{2}r)$  then
23:        $||(x, y) - (i, j)|| = r$  is a circle
24:     end if
25:   end for
26: end for

```

chaque boucle j . Une valeur d'accumulation de l'histogramme H_x est ainsi calculée à chaque fois qu'une colonne est lue et une valeur d'accumulation de l'histogramme H_y est calculée à chaque fois qu'une ligne est lue.

Cette modification a un double impact :

- un sur la latence de calcul parce que les maximums locaux de l'histogramme H_x et H_y peuvent être calculés dès qu'une colonne et qu'une ligne sont respectivement lues ;
- un sur l'utilisation des ressources parce que nous n'avons plus besoin de mémoriser les histogrammes H_x et H_y .

Pour valider le fonctionnement de notre algorithme [1DHT](#) modifié, nous avons fait les évaluations suivantes :

- dans une première évaluation, nous avons évalué l'algorithme pour différentes valeurs de f et pour une image synthétique avec des formes fermées représentées dans la figure 7.1. f correspond au *seuil de sensibilité* présent à la ligne 22 de l'algorithme 6. Le cercle vert indique qu'une figure fermée a été détecté. Comme cela est visible dans la figure 7.1, cet algorithme peut détecter différentes figures fermées. Nous constatons que la variation du seuil f permet de détecter différents types de figures fermées.

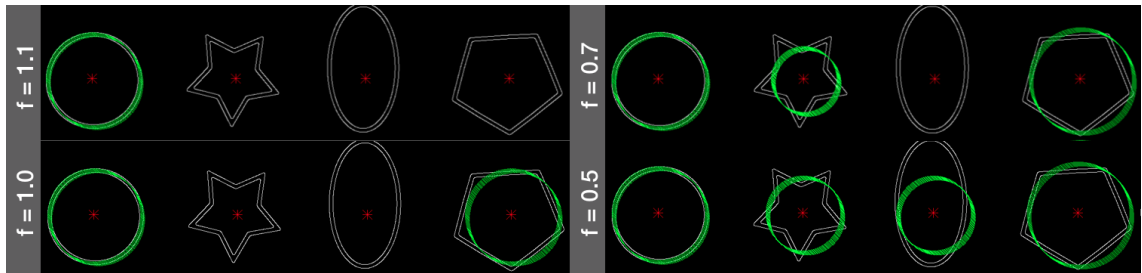


FIGURE 7.1 : Validation de l'algorithme avec différentes valeurs du seuil f

- Dans une deuxième évaluation, nous avons évalué la capacité de l'algorithme dans la détection des contours circulaires sur des images synthétiques. Pour cette évaluation nous avons créé des images synthétiques : 3000 images sans bruit contenant des cercles non superposés, un échantillon est visible sur la figure 7.2). 3000 images sans bruit contenant des cercles superposés dont un échantillon est visible sur la figure 7.3). 3000 images avec bruit contenant des cercles non superposés dont un échantillon est visible sur la figure 7.4). Les types de bruit utilisés sont un bruit Gaussien, un bruit flou, un bruit poivre

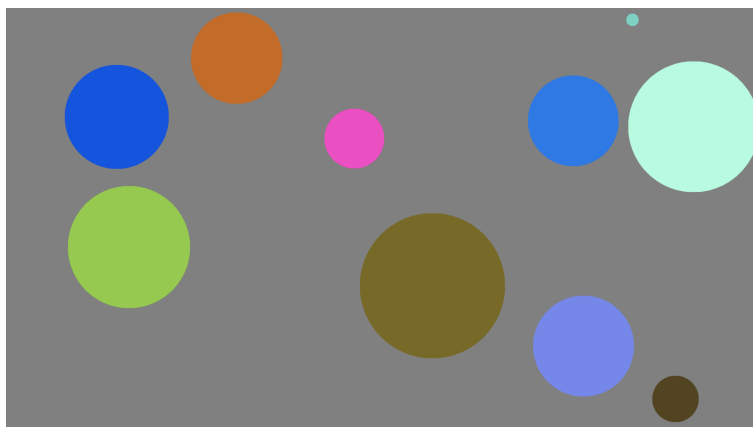


FIGURE 7.2 : Image synthétique sans bruit avec 10 cercles non superposés.

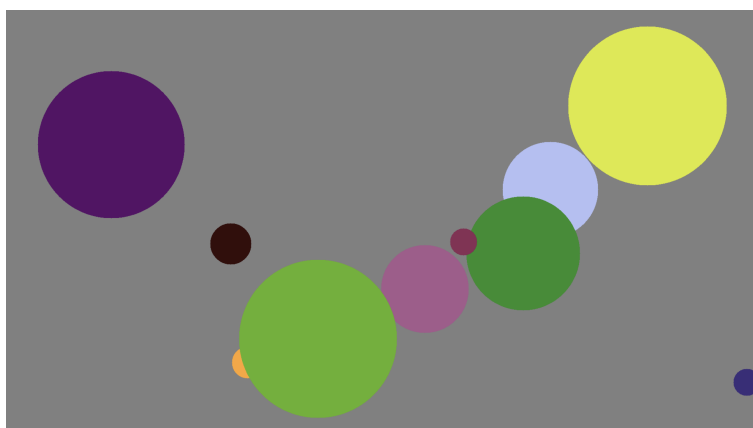


FIGURE 7.3 : Image synthétique sans bruit avec 10 cercles superposés.

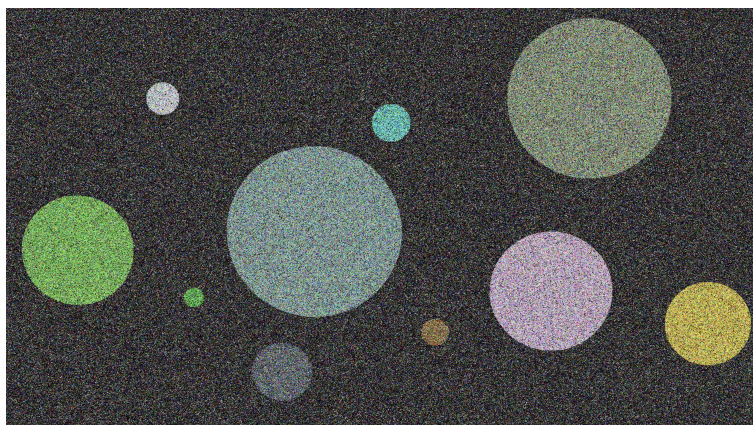


FIGURE 7.4 : Image synthétique avec bruit avec 10 cercles non superposés.

et sel avec une degré de PSNR ¹ de 8,44 et une rapport signal sur bruit de 13,06 dB. Chaque image contient 10 cercles positionnés aléatoirement. Nous avons comparé notre algorithme avec l'algorithme de la transformée de Hough original mis en œuvre dans la librairie OpenCv[95]. Les résultats de cette évaluation sont visibles dans le tableau 7.1. Notre algorithme a une performance similaire pour la détection des contours circulaires sur des images synthétiques sans bruit avec de cercles superposés et non superposés. Pour le cas d'images avec bruit notre algorithme démontre une performance inférieure à celle de la transformée de Hough de la librairie OpenCv.

TABLE 7.1 : Comparaison entre notre algorithme et la référence de l'algorithme de la transformée de Hough de la librairie OpenCv.

	notre transformée de Hough			transformée de Hough d'OpenCv		
1000 Images	sans bruit	sans bruit	avec bruit	sans bruit	sans bruit	avec bruit
Avec 10 cercles par image	non superposés	superposés	non superposés	non superposés	superposés	non superposés
Cercles détectés	99,3%	87,8%	73,6%	99,0%	91,3%	81,7%

Les résultats de la comparaison entre notre algorithme et celui mis en œuvre dans la librairie OpenCv montrent que :

- notre algorithme peut détecter des cercles avec un contour irrégulier même en présence de bruit.
- notre algorithme obtient des performances similaires à la version originelle de la transformée de Hough de la librairie OpenCv dans la détection des contours circulaires sur des images synthétiques, ce qui permet de valider le fonctionnement de notre algorithme.

Nous avons proposé un algorithme fonctionnel et conçu pour être intégré dans une [CE](#). Nous décrivons, maintenant, l'architecture que nous proposons pour la mise en œuvre de cet algorithme.

7.2 Architecture pour la mise en œuvre de l'algorithme 1DHT modifié

Le synoptique général de l'architecture que nous proposons est visible à la figure 7.5. Cette architecture s'appelle 1DHT IP, elle est composée de cinq modules que nous décrivons ci-dessous :

¹Peak Signal to Noise Ratio

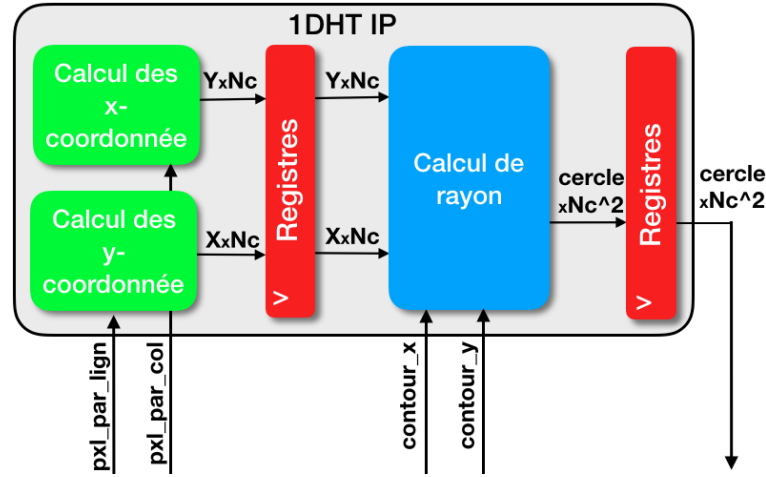


FIGURE 7.5 : Synoptique de l'architecture proposée 1DHT IP.

- les modules **calcul des x-coordonnées**, **calcul des y-coordonnées** : ces deux modules effectuent le calcul de N_c coordonnées en x et de N_c coordonnées en y, qui correspondent aux coordonnées possibles des centres des cercles, ou ellipses, dans l'image. N_c^2 centres candidats sont produits à partir de la combinaison de chaque coordonnée en x et chaque coordonnée en y ;
- le module **calcul du rayon** : ce module construit un histogramme pour chaque centre candidat en calculant la distance euclidienne entre le centre lui-même et chaque point du contour. La valeur de cette distance est accumulée dans son propre histogramme. Une fois l'histogramme totalement construit, ce module attribue la valeur de la distance euclidienne la plus accumulée comme le rayon de ce centre candidat. Si cette valeur est plus grande que $f * 4\sqrt{2}r$, le centre candidat et son rayon sont considérés comme étant un vrai cercle ;
- les modules **registres**, ces deux modules permettent de mémoriser les N_c coordonnées en x, les N_c coordonnées en y et les N_c^2 centres candidats.

Dans les paragraphes suivants nous allons décrire chaque module de notre architecture.

7.2.1 Calcul des coordonnées en x

Le module **calcul des x-coordonnées** est complètement pipeliné sur trois étages, le **calcul de l'histogramme**, le **filtrage des maxima locaux** et la **sélection de N_c coordonnées en x**. Il est représenté par la figure 7.6. Il calcule N_c coordonnées en x pour un diamètre maximum de D_m , avec $D_m > 1$.

Dans le premier étage, pour une image de taille $W * H$ chaque colonne de H bits est mémorisée dans une mémoire FIFO de taille $1 * H$ bits. Il faut un total une

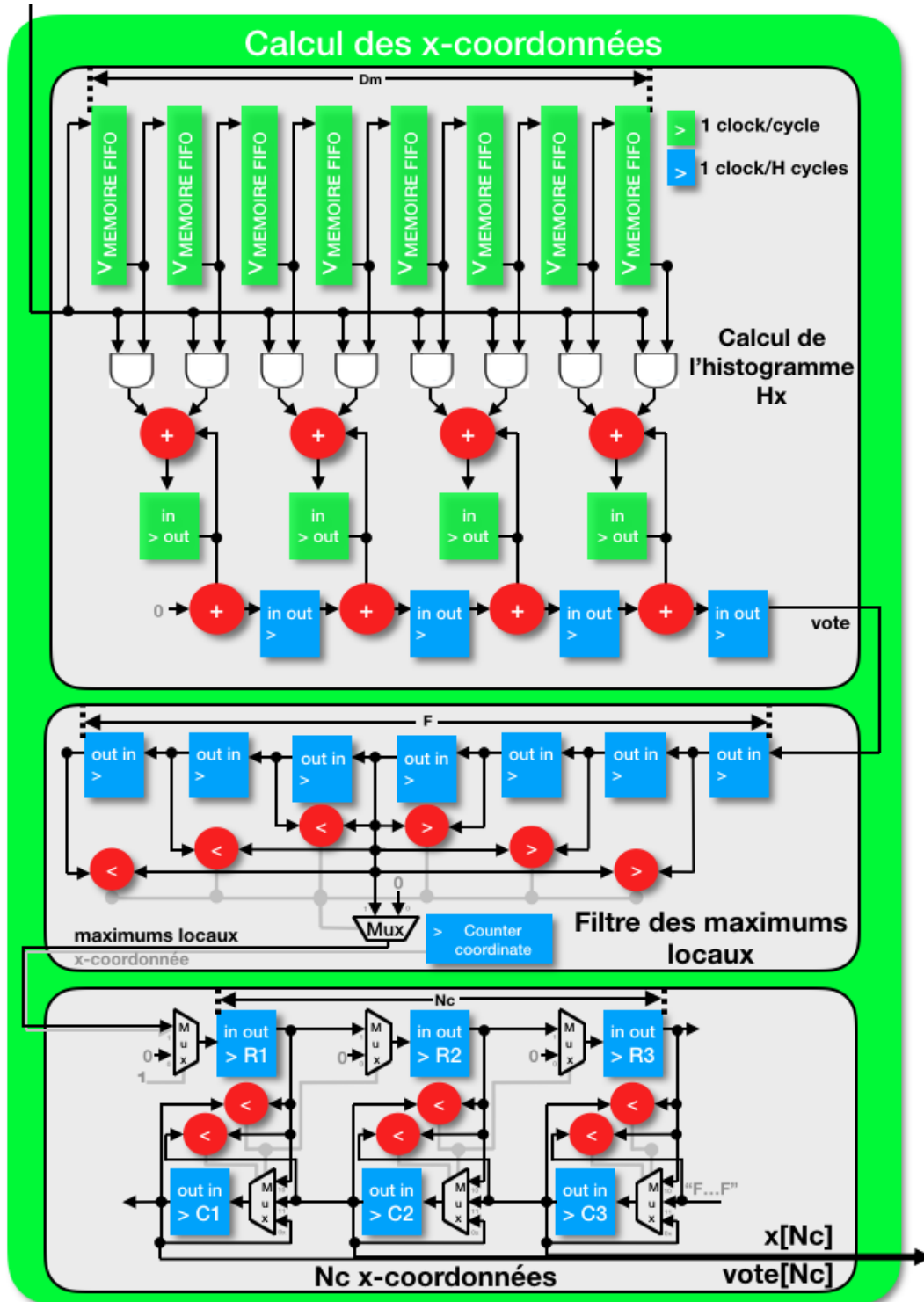


FIGURE 7.6 : Architecture pour le calcul des coordonnées en x.

mémoire de $D_m * H$ bits pour déplacer D_m colonnes, de cette manière il est possible d'accumuler les votes entre une colonne d'entrée et les D_m colonnes précédentes.

Il faut deux additionneurs et un registre pour réaliser la somme des votes entre la colonne d'entrée et les deux colonnes précédentes comme nous pouvons le voir dans la figure 7.6. Au total, il faut D_m additionneurs pour réaliser la somme des votes entre la colonne d'entrée avec les D_m colonnes précédentes. Le vote total est accumulé dans les registres, les blocs **bleus** dans la figure 7.6, pour être déplacés chaque H cycles d'horloge une fois que la colonne d'entrée a été complètement lue. Cette architecture calcule une valeur d'accumulation de l'histogramme H_x chaque fois qu'une colonne est lue, ces valeurs correspondent aux votes pour chaque coordonnée x .

Dans le second étage, les votes sont filtrés par une fenêtre glissante avec l'objectif de trouver les maximums locaux. La fenêtre glissante est de taille F et est utilisée pour comparer chaque vote avec les $\frac{F}{2}$ votes précédents et $\frac{F}{2}$ votes suivants. Ici F correspond à la distance euclidienne minimale entre deux centres. Les votes à la sortie de cet étage correspondent aux valeurs des coordonnées x les plus votées de l'histogramme soient les maxima locaux. Si dans l'image il y a deux cercles avec des centres séparés par une distance euclidienne plus petite que F pixels, seul le cercle avec le plus de points de contour sera détecté c'est à dire le plus grand cercle.

Enfin, dans le troisième étage, les maxima locaux sont stockés dans les registres R_i and C_i . Les registres R_i stockent les maxima locaux un par un en éliminant les valeurs répétées et les registres C_i stockent les N_c maxima locaux les plus grands. Les registres sont transférés en respectant les règles suivantes :

- Si $R_i > C_i$, alors $R_{i+1} = R_i$ sinon $R_{i+1} = 0$;
- Si $R_i > C_{i+1}$ et $R_i > C_i$, alors $C_i = C_{i+1}$
sinon, si $R_i > C_{i+1}$ et $R_i < C_i$ alors $C_i = R_i$ sinon $C_i = C_i$.

Dans cette architecture, les valeurs les plus grandes seront progressivement transférées du côté droit à travers les registres C_i . Ce processus est exécuté jusqu'à ce que toutes les colonnes soient lues. Finalement les plus grands maxima locaux sont stockés dans les registres C_i et leurs coordonnées respectives correspondent aux N_c coordonnées en x des cercles candidats.

N_c coordonnées en y sont calculées avec la même architecture dans le module **calcul des y-coordonnées**.

Le nombre de ressources total et la latence pour le calcul des coordonnées en x et des coordonnées en y sont listés dans le tableau 7.2.

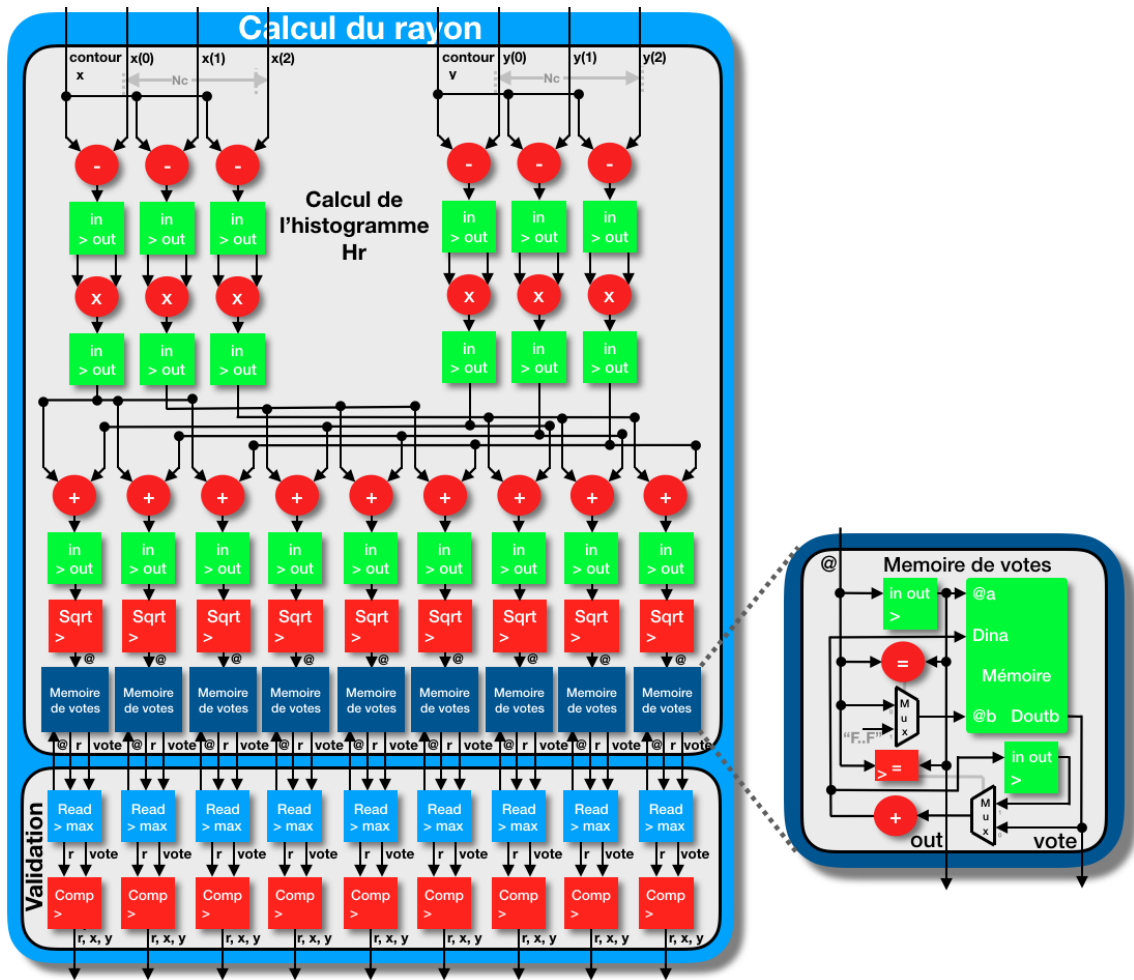


FIGURE 7.7 : Architecture pour le calcul de rayon.

TABLE 7.2 : Calcul des coordonnées en x et des coordonnées en y pour une image de taille $H * W$.

Module	Ressources nécessaires				Latence [cycles]
	Mémoire[bits]	Additionneurs	Registres	Comparateurs	
Calcul des x	$D_m * H$	D_m	$D_m + F + 2N_c$	$F + 2N_c - 1$	$W * H + \frac{D_m}{2} + \frac{F}{2} + N_c + 2$
Calcul des y	$D_m * H$	D_m	$D_m + F + 2N_c$	$F + 2N_c - 1$	$W * H + \frac{D_m}{2} + \frac{F}{2} + N_c + 2$
Total : Calcul des x et y	$D_m(H + W)$	$2D_m$	$2D_m + 2F + 4N_c$	$2F + 4N_c - 2$	$W * H + \frac{D_m}{2} + \frac{F}{2} + N_c + 2$

7.2.2 Calcul du rayon

A la suite des calculs provenant des modules **calcul des x-coordonnées** et **calcul des y-coordonnées**, nous obtenons N_c^2 centres candidats pour lesquels il faut calculer un rayon.

Nous avons conçu une architecture pipelinée composée de deux étages pour le calcul du rayon des centres candidats, elle est visible la figure 7.7.

Dans le premier étage, l'histogramme H_r est construit pour chaque cercle candidat en parallèle. Cette architecture calcule la distance euclidienne entre chaque centre candidat et tous les point de contours. Nous utilisons l'algorithme de Cordic [92] pour réaliser le calcul de la racine carrée codée avec des entiers sur 16bits. Le temps de calcul est de 8 cycles d'horloge.

Pour stocker les votes des distances euclidiennes nous avons réalisé une architecture, une mémoire spécifique représenté dans la figure 7.7. Dans cette architecture, chaque bloc mémoire permet une lecture et une écriture simultanée en un seul cycle, en évitant les collisions.

Une fois que tous les points de l'image sont lus, dans le second étage, pour chaque centre candidat, toutes les valeurs de la mémoire sont lues afin de trouver la valeur de la distance euclidienne la plus votée, cette lecture s'effectue en $\frac{D_m}{2}$ cycles. Cette valeur représente le rayon r pour ce centre candidat. Nous comparons ce rayon r avec le seuil $4\sqrt{2}r * f$ défini dans [57] pour déterminer s'il correspond à un cercle. Il est possible de modifier ce seuil pour faire varier la sensibilité de détection.

Chaque rayon qui correspond à un cercle et son centre respectif sont stockés dans le registre de stockage montré dans la figure 7.5.

Le nombre de ressources et la latence pour le calcul des rayons de N_c^2 centres candidats sont listés dans le tableau 7.5. Les rayons sont codés avec L bits et les votes sont codés avec M bits.

TABLE 7.3 : Calcul des rayons pour une image de taille $H \times W$

Module	Ressources nécessaires				Latence [cycles]
	Mémoire [bits]	Opérateurs (+, -, *)	Registres	Comparateurs	
Calcul des rayons	$N_c^2 * 2^L M$	$4N_c + 2N_c^2$	$4N_c + 3N_c^2$	$4N_c^2$	$W * H + 18 + \frac{D_m}{2}$

7.3 Mise en œuvre de l'architecture du 1DHT IP sur un FPGA

Dans cette section nous décrivons la mise en œuvre des différents modules de notre architecture 1DHT IP dans un FPGA. Pour notre cas d'étude nous avons choisi le FPGA Xilinx Artix-7 qui a des LUT², des opérateurs DSP, des mémoires BRAM de 18Kbits et 36Kbits et qui est celui mis en place dans la partie logique programmable du circuit Zynq.

7.3.1 Modules Calcul des x-coordonnées et Calcul des y-coordonnées

La mise en œuvre de ces modules sur un FPGA est visible sur la figure 7.8 dans laquelle est représenté le module **Calcul des x-coordonnées** jumeau du module **Calcul des y-coordonnées**.

Des BRAM ont été utilisées pour la mise en œuvre des mémoires FIFO. Comme pour une image de taille $H * W$, une BRAM de 36Kb peut déplacer $\lceil 36K * 2^{-\lceil \log_2 H \rceil} \rceil$ colonnes, il faut un total de $\frac{D_m}{\lceil 36K * 2^{-\lceil \log_2 H \rceil} \rceil}$ BRAM pour déplacer D_m colonnes.

Les deux additionneurs et le registre pour réaliser la somme des votes entre la colonne d'entrée et les deux colonnes précédentes ont été mis en œuvre avec un seul bloc DSP comme nous pouvons le voir dans la figure 7.8. En total, il faut $\frac{D_m}{2}$ DSP pour réaliser la somme des votes entre la colonne d'entrée avec les D_m colonnes précédentes.

Les opérateurs logiques, les comparateurs et les registres ont été mis en œuvre à l'aide de LUT.

Le nombre de ressources et la latence pour les modules **calcul des x-coordonnées** et **calcul des y-coordonnées** dans un FPGA sont indiqués dans le tableau 7.2.

7.3.2 Module Calcul du rayon

La mise en œuvre de ce module dans un FPGA est visible sur la figure 7.11.

²Look Up Table

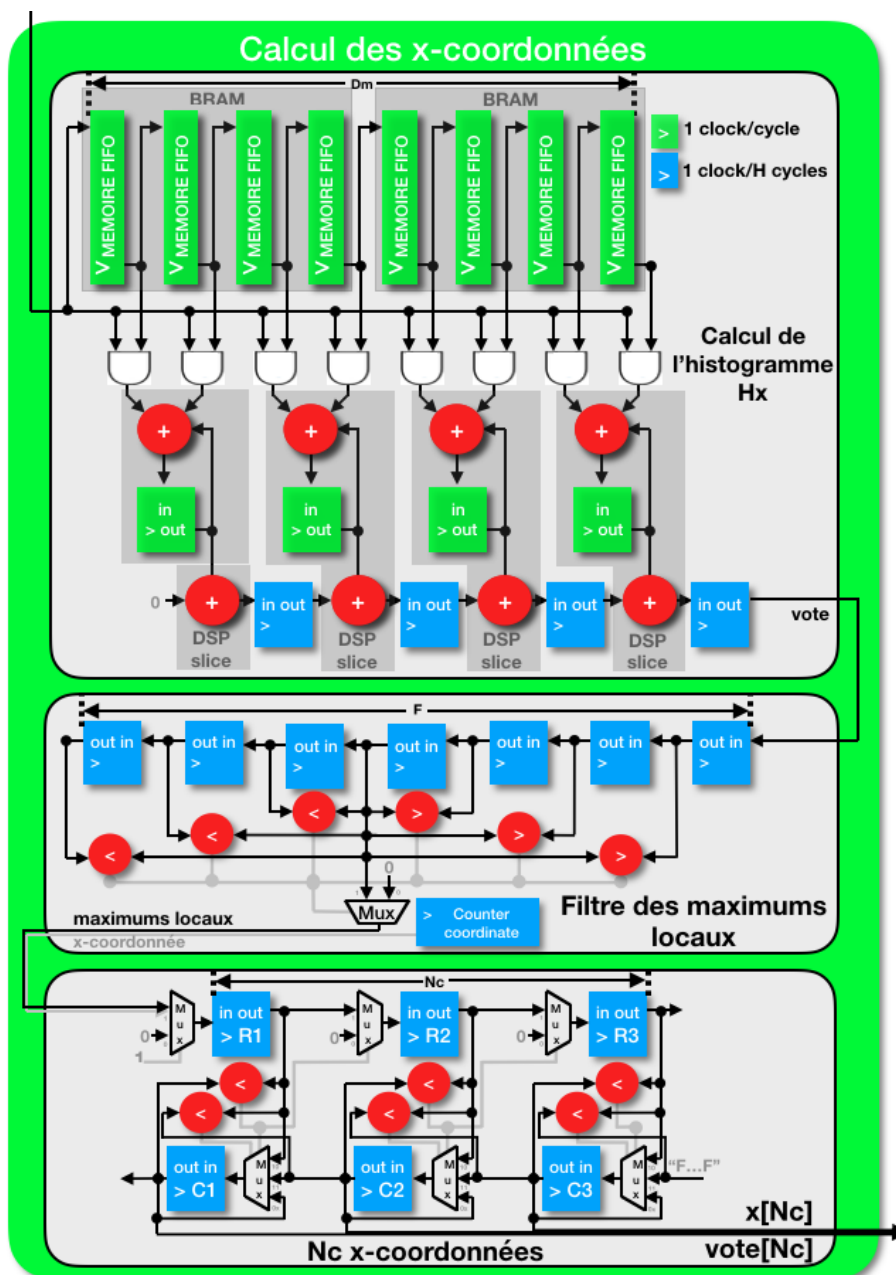


FIGURE 7.8 : Architecture FPGA pour le module calcul des x-coordonnées.

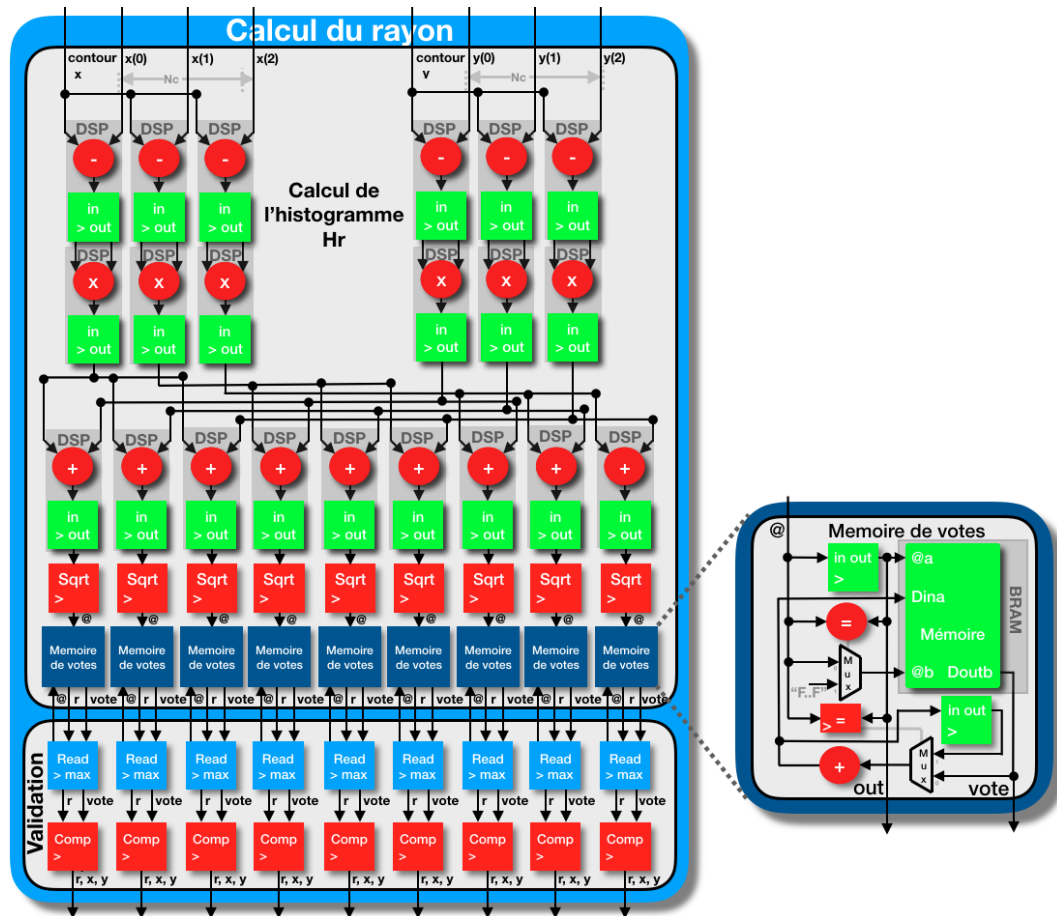


FIGURE 7.9 : Architecture FPGA pour le calcul de rayon.

TABLE 7.4 : Ressources des modules **Calcul des x-coordonnées** et **Calcul des y-coordonnées** pour une image de taille $H * W$.

Module	Ressources nécessaires		Latence [cycles]
	BRAM de 36 Kbits	Blocs DSP	
Calcul des x- et y-coordonnés	$\frac{D_m}{\lceil 36K * 2^{-\lceil \log_2 H \rceil} \rceil} + \frac{D_m}{\lceil 36K * 2^{-\lceil \log_2 W \rceil} \rceil}$	D_m	$W * H + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$

Chaque somme ou soustraction est exécutée en 2 cycles, chaque multiplication est exécutée en 3 cycles et chaque opération est implantée avec un bloc DSP. Nous utilisons l'IP "Xilinx CORDIC IP core" [92] pour mettre en œuvre l'algorithme CORDIC pour le calcul de la racine carrée codée avec des entiers sur 16bits, le temps de calcul est de 8 cycles d'horloge.

Pour la mise en œuvre de la mémoire de votes, chaque bloc mémoire utilise une BRAM de 18 Kbits qui permet une lecture et une écriture simultanée en un seul cycle.

TABLE 7.5 : Calcul des rayons pour une image de taille $H \times W$

Module	Ressources nécessaires		Latence [cycles]
	BRAM de 36 Kbits	blocs DSP	
Calcul des rayons	$\frac{N_c^2}{2}$	$4N_c + N_c^2$	$W * H + 18 + \frac{D_m}{2}$

Au total, il faut $4N_c + N_c^2$ blocs DSP et $\frac{N_c^2}{2}$ BRAM de 36 Kbits pour réaliser ce calcul de vote pour N_c^2 centres candidats. Le calcul du rayon est réalisé en $W * H + 18 + \frac{D_m}{2}$ cycles. Ces chiffres sont visibles dans le tableau 7.5.

7.4 Mise en œuvre de l'architecture du 1DHT IP dans un Système sur Puce

Pour valider notre architecture, nous l'avons intégrée dans un système sur puce, le Zynq-7000 XC7Z020 en utilisant la carte de développement Digilent ZedBoard, pour valider fonctionnellement l'architecture et réaliser des mesures du temps d'exécution.

Le Zynq n'est pas notre plateforme finale, il ne peut pas être intégré dans une CE, c'est notre plateforme de test et de validation pour des questions de facilité de mise en œuvre et parce que la partie reconfigurable est similaire à celle que nous envisageons d'utiliser, un Spartan 7 de dimension $8 * 8 \text{ mm}^2$.

Dans la figure 7.10 nous illustrons l'intégration de notre architecture 1DHT IP sur

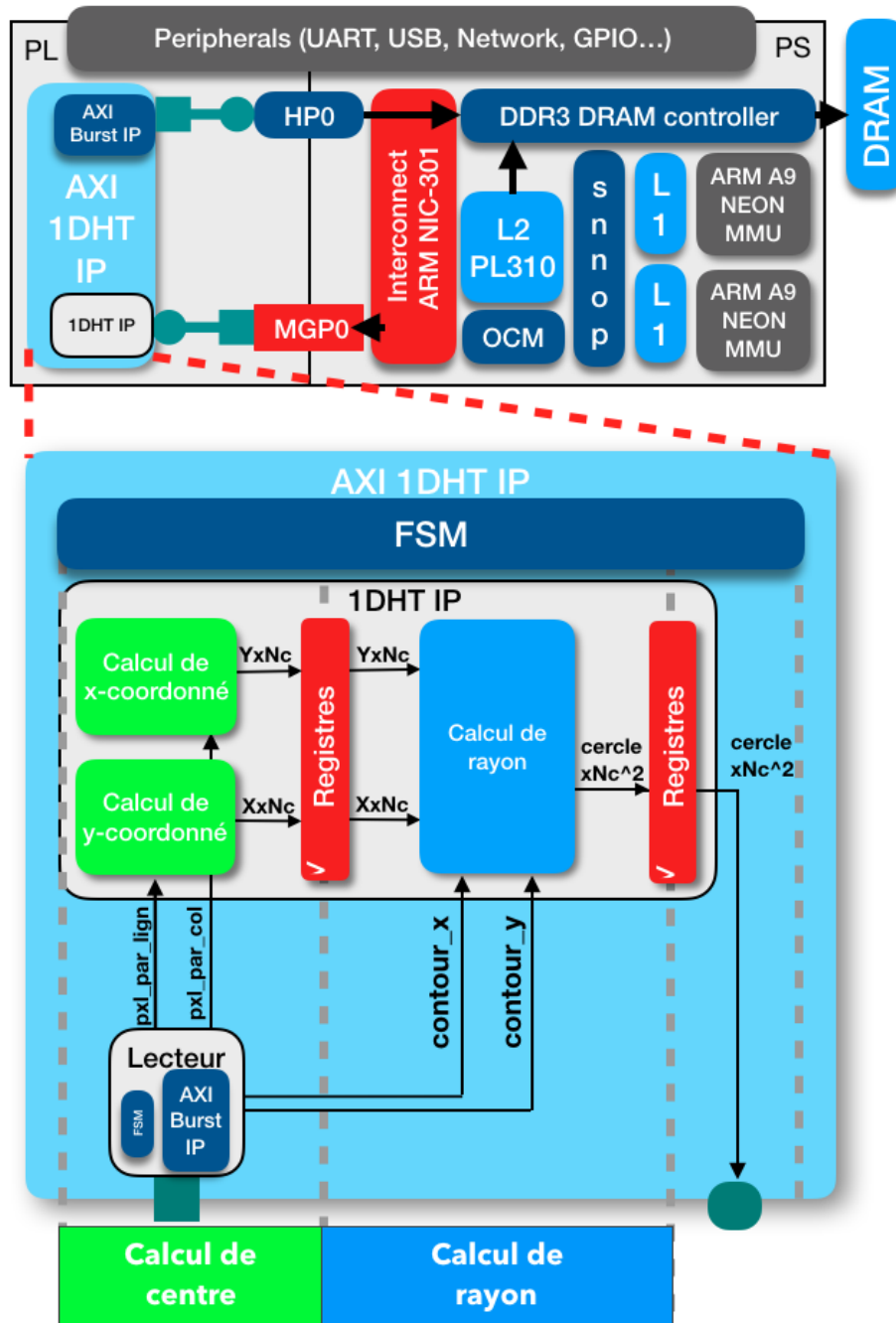


FIGURE 7.10 : Diagramme de l'IP 1DHT intégré dans un système Zynq.

un Zynq. Nous utilisons un module de communication AXI AMBA pour interfacer la partie reconfigurable avec la partie [ARM](#) Cortex A9 du Zynq. Ce module utilise un port HP AXI maître et un port AXI esclave pour lire les images et envoyer les résultats des calculs des cercles détectés. L'IP intégrant les modules de calcul et le module de communication s'appelle AXI 1DHT IP comme nous pouvons le voir sur la figure 7.10. Pour lire et écrire l'image dans la DRAM nous utilisons la technique de double buffers : le processeur écrit à une adresse en mémoire pendant que l'accélérateur lit une autre adresse de la mémoire.

Comme notre architecture a deux étapes de calcul, nous utilisons quatre adresses (@1, @2, @3, @4) contigües de la DRAM afin de lire et d'écrire les images en même temps.

Pour cette mise en œuvre, nous avons réalisé un pipeline sur trois niveaux :

1. niveau 1 : le processeur [ARM](#) CortexA9 écrit l'image aux adresses @1 et @2 de la mémoire DRAM et reçoit les cercles envoyées par l'IP AXI 1DHT.
2. Niveau 2 : l'IP AXI 1DHT lit l'image à l'adresse @1 et calcule les centres candidats.
3. Niveau 3 : l'IP AXI 1DHT lit l'image à l'adresse @2 et calcule les rayons de chaque cercle candidat.

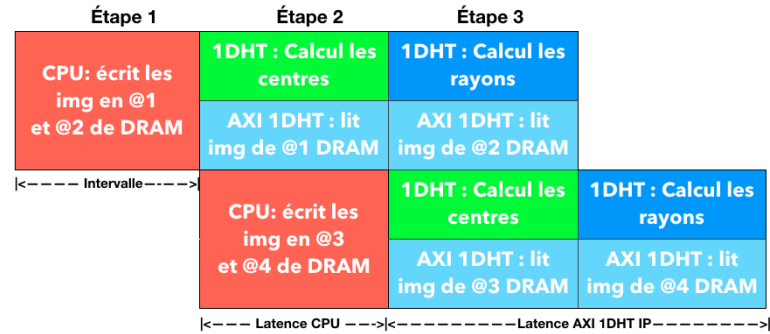


FIGURE 7.11 : Intégration de l'IP "AXI 1DHT" sur Zynq : pipeline au niveau système

Ce pipeline est visible dans la figure 7.11.

Pour le calcul des centres candidats il faut $W * H + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$ cycles et pour le calcul des rayons il faut $W * H + 18 + \frac{D_m}{2}$ cycles. La latence totale de l'IP "AXI 1DHT IP" correspond à deux fois le temps de calcul des centres candidats qui est le chemin critique et donc la plus grande latence du pipeline.

Finalement pour mettre en œuvre l'IP "AXI 1DHT IP" il faut $\frac{D_m}{\lceil 36K * 2^{-E(\log_2 H)} \rceil} + \frac{D_m}{\lceil 36K * 2^{-E(\log_2 W)} \rceil} + \frac{N_c^2}{2}$ BRAM de 36Kbits et $D_m + 4N_c + N_c^2$ blocs DSP . Cela permet de détecter N_c^2 cercles candidats avec un diamètre maximal de D_m et séparés de minimum F pixels dans une image de taille $W * H$. Il faut $2(W * H + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c)$

TABLE 7.6 : Caractéristiques de mise en œuvre de l'IP "AXI 1DHT IP" pour une image de taille $H \times W$

Module	Ressources nécessaires		Latence [cycles]
	BRAM de 36 Kbits	blocs DSP	
Calcul des x- et y-coordonnés	$\frac{D_m}{\lceil 36K * 2^{-E(\log_2 H)} \rceil} + \frac{D_m}{\lceil 36K * 2^{-E(\log_2 W)} \rceil}$	D_m	$W * H + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$
Calcul des rayons	$\frac{N_c^2}{2}$	$4N_c + N_c^2$	$W * H + 18 + \frac{D_m}{2}$
Estimation pour 1920x1080 pixels.	24.5	153	4147358

cycles de latence et $W * H + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$ cycles d'intervalle entre deux images pour réaliser ce calcul, voir tableau 7.6.

Ce modèle analytique des performances de notre IP, montre que la latence et l'intervalle entre deux images sont directement liés à la taille de l'image. Le goulot d'étranglement en augmentant la résolution de l'image provient aussi du nombre limité de BRAM disponibles sur le Zynq. Le diamètre maximum des cercles D_m et le nombre de cercles que ce module peut détecter se retrouvent limités par le nombre de BRAM et de blocs DSP disponibles.

Nous avons réalisé une estimation pour détecter $N_c^2 = 25$ cercles avec un diamètre maximal de $D_m = 108$ pixels séparés d'une distance minimales de $F = 40$ pixels pour des images de taille de 1920x1080 pixels, les résultats sont visibles dans le tableau 7.6.

TABLE 7.7 : Mise en œuvre de l'IP "AXI 1DHT IP" sur le Zynq.

Architecture	1DHT IP	AXI CHT IP IP core	Zynq Système	Utilisation [%]
LUT	20637	20770	21638	38,5%
FF	18629	18697	19706	18,5%
BRAM (36Kb)	24,5	24,5	24,5	17,5%
DSP	153	153	153	69,5%
Freq. maximum[MHz]	149,16	147,99	128,89	-
Latence[ms]	27,81	28,03	32,18	-
Intervalle[ms]	13,90	14,01	16,09	-
Cadence[Mpixel/s]	149,15	147,98	128,89	-
fps[frame/s]	72	71	62	-

Nous avons ensuite réalisé la mise en œuvre sur le Zynq et l'exécution de l'IP pour détecter $N_c^2 = 25$ cercles avec un diamètre maximal de $D_m = 108$ pixels séparés d'une distance minimale de $F = 40$ pixels pour des images de taille de 1920x1080 pixels. Nous avons utilisé 11 bits pour coder les coordonnées x et y , 8 bits pour coder les rayons r et 20 bits pour coder les votes. Les détails de cette mise en œuvre sont listés dans la table 7.7.

A la lecture du tableau 7.7 nous remarquons que notre transformée de Hough peut fonctionner jusqu'à une fréquence de 149,16 MHz lorsqu'elle fonctionne seule. Quand elle est intégrée dans le système global elle peut fonctionner jusqu'à une fréquence de 128,885 MHz.

Notre architecture peut traiter une image de taille 1920x1080 pixels en moins de 40ms avec une cadence de 62 images per seconde.

Si nous appliquons une fréquence de 128,89MHz au modèle analytique que nous avons réalisé et qui est décrit dans le tableau 7.6, nous obtenons une latence de 32,18ms, une utilisation de 24,5 BRAMS et de 153 DSP. Ce résultat est en concordance avec ceux obtenus dans le tableau 7.7 ce qui valide le modèle analytique que nous avons réalisé.

Nous avons comparé notre IP "1DHT" avec l'état de l'art montré dans le tableau 6.1. Pour cela nous avons estimé le temps d'exécution, le nombre de BRAM et DSP de notre architecture pour différentes tailles d'images afin de la mettre à l'échelle des différents travaux de l'état de l'art. Ces estimations sont visibles dans le tableau 7.8. Nous avons réalisé la comparaison dans le tableau 7.9. Nous remarquons que notre architecture offre une vitesse de traitement supérieure à toutes les architectures de l'état de l'art.

TABLE 7.8 : Notre méthode pour différentes tailles d'image.

Notre méthode	Artix7 avec une fréquence de 128.89MHz					
	Estimée et simulée					Mise en œuvre
Résolution[pixel]	320x240	400x400	512x512	640x480	1920x1080	1920x1080
BRAM of 36Kb	15,5	16,5	16,5	17,5	24,5	24,5
DSP48 slices	153	153	153	153	153	153
Latence [ms]	1,19	2,48	4,07	4,77	32,18	32,18
Intervalle [ms]	0,60	1,24	2,03	2,38	16,09	16,08
fps[frame/s]	1676	805	491	419	62	62
Cadence -[Mpixel/s]	128,8	128,9	128,9	128,9	128,9	128,9

TABLE 7.9 : Comparaison de cadence de notre méthode avec l'état de l'art.

Auteurs / Résolution	Latence[ms]/freq[Hz]		Cadence[$\frac{Mpxl}{s}$]		Mémoire[Kb]	
	Auteurs	nous	Auteurs	nous	Auteurs	nous
[38] / 800x600	33,3 / -	7,5 / 128,9	14,40	128,9	256	666
[100] / 400x400	5,3/181,8	2,48 / 128,9	29,97	128,9	5562	594
[83] / 640x480	10/250	4,77 / 128,9	30,72	128,9	-	630
[52] / 512x512	7,81/135	4,07 / 128,9	33,55	128,9	512	594
[58] / 320x240	3,46/203	1,19 / 128,9	22,20	128,8	1872	558

Ces résultats sont positifs en terme de latence puisque notre architecture satisfait la contrainte de 40 ms par image et qu'elle est à l'état de l'art sur ce point. Mais

ils sont pour l'instant des résultats issus de la mise en œuvre sur Zynq, il nous faut aller plus loin et estimer les performances sur un FPGA compatible avec une intégration dans une CE. Pour cela nous avons choisi comme cible le FPGA Spartan 7 XC7S15-CPGA196 de chez Xilinx. Ce circuit avec son boîtier, possède une surface de $8 \times 8 \text{ mm}^2$, ce boîtier a une diagonale de 11,3 mm qui permet une intégration dans une CE de diamètre de 11,6mm.

Les résultats de la mise en œuvre sont présentés dans le tableau 7.10.

TABLE 7.10 : Résultats de la mise en œuvre de l'IP 1DHT sur le Spartan 7 XC7S15.

Ressources	Disponibles	Utilisées	Utilisées %
LUT	8000	20637	258%
LUTRAM	2400	275	11%
FF	16000	18629	116%
BRAM36K	10	24,5	245%
DSP	20	153	765%

À la lecture du tableau 7.10, nous pouvons remarquer que les ressources sont insuffisantes pour garantir l'intégration de l'IP 1DHT. Nous avons réalisé une analyse plus poussée des ressources nécessaires pour chaque module de l'IP, les résultats sont indiqués dans le tableau 7.11.

TABLE 7.11 : Ressources nécessaires pour la mise en œuvre de l'IP 1DHT

Calcul	BRAM36K	DSP	LUT	FF
Calcul des centres candidats	12	108	7656	6544
x-coordonnées	6	54	3828	3272
y-coordonnées	6	54	3828	3272
Calcul de rayons	12,5	45	7751	6471

Nous remarquons que la partie qui utilise le plus de BRAM et de DSP est le calcul des centres candidats.

À l'aide de ce résultat, nous avons choisi d'optimiser ce module pour permettre l'embarquabilité dans la CE de la transformée de Hough.

Nous présentons cette optimisation dans le chapitre suivant.

Chapitre 8

Optimisation de l'IP 1DHT en surface

Sometimes it is the people no one imagines anything of who do the things that no one can imagine.

Alan Turing

Nous avons analysé le calcul des centres candidats de l'algorithme [1DHT](#) proposé par Goneid et al. [45] avec pour objectif d'optimiser la surface de notre architecture. L'ambition est de pouvoir effectuer les calculs en temps réel en moins de 40 ms et avec une surface compatible pour une intégration au sein d'une [CE](#), c'est à dire une surface maximale de $8 * 8mm^2$ et ceci pour des images de taille $1080 * 1920$ pixels.

8.1 Optimisation du calcul des centres candidats

L'algorithme de Goneid est constitué principalement de deux étapes : le calcul des centres candidats et le calcul du rayon pour chaque centre candidat.

Nous avons proposé un algorithme équivalent pour l'accélération du calcul des coordonnées en x des centres candidats, voir l'algorithme 6. Nous avons proposé un algorithme similaire pour l'accélération du calcul des coordonnées en y des centres candidats.

Algorithm 7 Algorithme 1DHT pour le calcul des coordonnées en x

```
1: for chaque colonne  $j$  de 1 à  $W - Dm$  do
2:   for chaque colonne  $k$  de  $j + 1$  à  $j + Dm$  do
3:     for chaque ligne  $i$  de 1 à  $H$  do
4:        $H_x(\frac{j+k}{2}) = H_x(\frac{j+k}{2}) + I(i, j) * I(i, k)$ 
5:     end for
6:   end for
7: end for
```

Pour obtenir l'algorithme 6, nous avons pris en compte tous les points de l'image au lieu de ne prendre que les points du contours de l'image. Nous avons aussi changé l'ordre des boucles. Nous avons proposé une architecture pour la mise en œuvre de cet algorithme où les boucles i et k ont été réalisées en parallèle. Dans cette architecture, pour déplacer des colonnes nous avons utilisé des mémoires FIFO.

Pour accumuler les votes entre une colonne d'entrée et deux autres colonnes, nous avons utilisé des additionneurs et des registres. Pour réaliser les mémoires FIFO nous avons utilisé une mémoire BRAM, pour les additionneurs nous avons utilisé un bloc DSP, et pour les registres nous avons eu recours à plusieurs bascules Flip-Flop.

L'avantage de notre méthode est que nous obtenons la valeur d'accumulation de l'histogramme H_x chaque fois qu'une colonne est lue.

Après l'analyse de l'algorithme 6, nous constatons que bien que le ré-ordonnement des boucles accélère le calcul, lors du déplacement des colonnes les ressources mobilisées sont importantes : 6 BRAM de 36Kb et 54 DSP, voir le tableau 7.11.

Il est nécessaire de trouver un autre moyen d'obtenir les valeurs d'accumulation de l'histogramme H_x sans utiliser un trop grand nombre de registres.

Pour résoudre ce problème nous proposons de remplacer le déplacement des colonnes par le déplacement de pixels et aussi de remplacer les opérations d'addition et de multiplication par de simples opérateurs logiques.

A partir de l'algorithme 6 pour le calcul des coordonnées en x , nous pouvons établir l'équation 8.1 :

$$H_x\left(\frac{j+k}{2}\right) = \sum_{j=1}^{W-Dm} \sum_{k=j+1}^{j+Dm} \sum_{i=1}^H I_{(i,j)} * I_{(i,k)} \quad (8.1)$$

Si nous ré-ordonnons les sommes nous obtenons l'équation 8.2 :

$$H_x\left(\frac{j+k}{2}\right) = \sum_{i=1}^H \sum_{j=1}^{W-Dm} I_{(i,j)} * \left(\sum_{k=j+1}^{j+Dm} I_{(i,k)} \right) \quad (8.2)$$

En réalisant un changement de variable $n = \frac{j+k}{2}$ nous obtenons l'équation 8.3 :

$$H_x(n) = \sum_{i=1}^H \sum_{j=1}^{W-Dm} I_{(i,j)} * \left(\sum_{n=j+1}^{j+\frac{Dm}{2}} (I_{(i,2n-j-1)} + I_{(i,2n-j)}) \right) \quad (8.3)$$

Un nouveau ré-ordonnement nous permet d'obtenir l'équation 8.4 :

$$H_x(n) = \sum_{i=1}^H \sum_{j=1}^{W-Dm} \sum_{n=j+1}^{j+\frac{Dm}{2}} (I_{(i,j)} * I_{(i,2n-j-1)} + I_{(i,j)} * I_{(i,2n-j)}) \quad (8.4)$$

Comme $I_{(i,j)}$ est un pixel binaire, nous pouvons remplacer la multiplication par un simple *ET* logique.

Pour optimiser la surface, nous pouvons approximer l'addition avec un *OU* logique, bien que nous perdons un vote à chaque fois qu'il y a deux contours consécutifs, cette perte est uniforme tout au long de l'image, cela veut dire que cette perte ne change pas la forme de l'histogramme, nous obtenons les mêmes maxima locaux. Finalement, l'équation 8.4 peut être écrite sous la forme de l'équation 8.5.

$$H_x(n) = \sum_{i=1}^H \sum_{j=1}^{W-Dm} \sum_{n=j+\frac{Dm}{2}}^{j+\frac{Dm}{2}+1} (I_{(i,j)} \text{ AND } I_{(i,2n-j-1)}) \text{ OR } (I_{(i,j)} \text{ AND } I_{(i,2n-j)}) \quad (8.5)$$

Il est possible maintenant de décrire l'équation 8.5 sous forme d'un algorithme qui permet d'obtenir le même histogramme H_x que précédemment, l'algorithme 8.

Algorithm 8 Algorithme 1DHT optimisé en surface pour le calcul des coordonnées en x

```

1: for chaque ligne  $i$  de 1 à  $H$  do
2:   for chaque colonne  $j$  de 1 à  $W - Dm$  do
3:     for chaque colonne  $n$  de  $j + 1$  à  $j + \frac{Dm}{2}$  do
4:        $H_x(n) = H_x(n) + (I_{(i,j)} \text{ AND } I_{(i,2n-j-1)}) \text{ OR } (I_{(i,j)} \text{ AND } I_{(i,2n-j)})$ 
5:     end for
6:   end for
7: end for
    
```

L'avantage de l'algorithme 8 est que nous obtenons la valeur d'accumulation de l'histogramme H_x pour chaque boucle de n , cela signifie que nous obtenons la valeur d'accumulation de l'histogramme H_x chaque fois qu'un pixel d'une ligne est lu.

Deux autres avantages apparaissent qui sont d'une part l'utilisation d'opérateurs booléens simples, les *ET* et *OU* logiques, et d'autre part pour les calculs des votes le décalage de pixels d'un bit au lieu de colonnes de W bits. Grâce à ces optimisations le calcul des votes peut être effectué par des LUT au lieu de DSP et le décalage par des bascules Flip-Flops au lieu des BRAM.

Le coût à payer pour cette optimisation est une mémoire pour stocker ces valeurs d'accumulation, ce qui a un impact dans la latence parce qu'il faut lire toute une image pour avoir l'histogramme complet H_x .

Pour valider le fonctionnement de notre algorithme [1DHT](#) optimisé en surface, nous avons évalué la capacité de notre algorithme dans la détection des polypes sur 2967 images [FHD](#) de colonoscopie. Un exemple de traitement de notre algorithme sur une image de coloscopie est visible sur la figure 8.1.

Ces images correspondent aux vidéos 11, 12, 13, 16, 18 et 19 de la base de données ASU-Mayo [89] qui ont une résolution de $1920 * 1080$.

Pour cette analyse, nous avons comparé notre algorithme avec l'algorithme de la transformée de Hough de la librairie OpenCv[95]. Les résultats sont listés dans

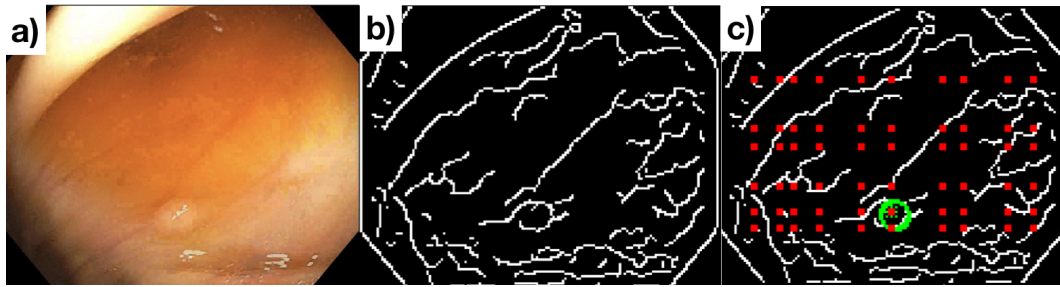


FIGURE 8.1 : a) image réel, b) image de contour et c) cercle détecté en vert.

le tableau 8.1. Les deux algorithmes extraient 70% des polypes au moins dans une ROI qui contient au moins 50% du polype.

TABLE 8.1 : Comparaison de notre algorithme optimisé en surface par rapport à la transformée de Hough de la librairie OpenCv

	Notre Transformée d'Hough	Transformée d'Hough OpenCv
Polypes extraits au moins dans une ROI	79,6%	88,9%
* Polypes extraits au moins dans une ROI qui contient au moins 90% du polype	50,5%	34,7%
* Polypes extraits au mois dans une ROI qui contient au moins 50% du polype	73,4%	71,1%
* Polypes extraits au mois dans une ROI qui contient au moins 10% du polype	79,6%	88,9%
Polypes n'apparaissant dans aucune ROI	20,4%	11,1%

À partir ces évaluations, nous pouvons remarquer que :

- notre algorithme peut détecter des cercles avec une contour irrégulier, comme celui des polypes, même en présence de bruit.
- notre algorithme a montré des performances similaires à la version originale de la transformée d'Hough de la librairie OpenCv sur des images de coloscopie, ce qui permet de valider le fonctionnement de notre algorithme.

Armé de ce nouvel algorithme nous avons réalisée une architecture qui le met en œuvre.

8.2 Architecture proposée pour la mise en œuvre du calcul des coordonnées en x

Pour mettre en œuvre l'algorithme 8, nous proposons une nouvelle architecture visible dans la figure 8.2. Dans cette architecture, les calculs correspondants à la boucle sur la variable n de l'algorithme 8 sont réalisés en parallèle.

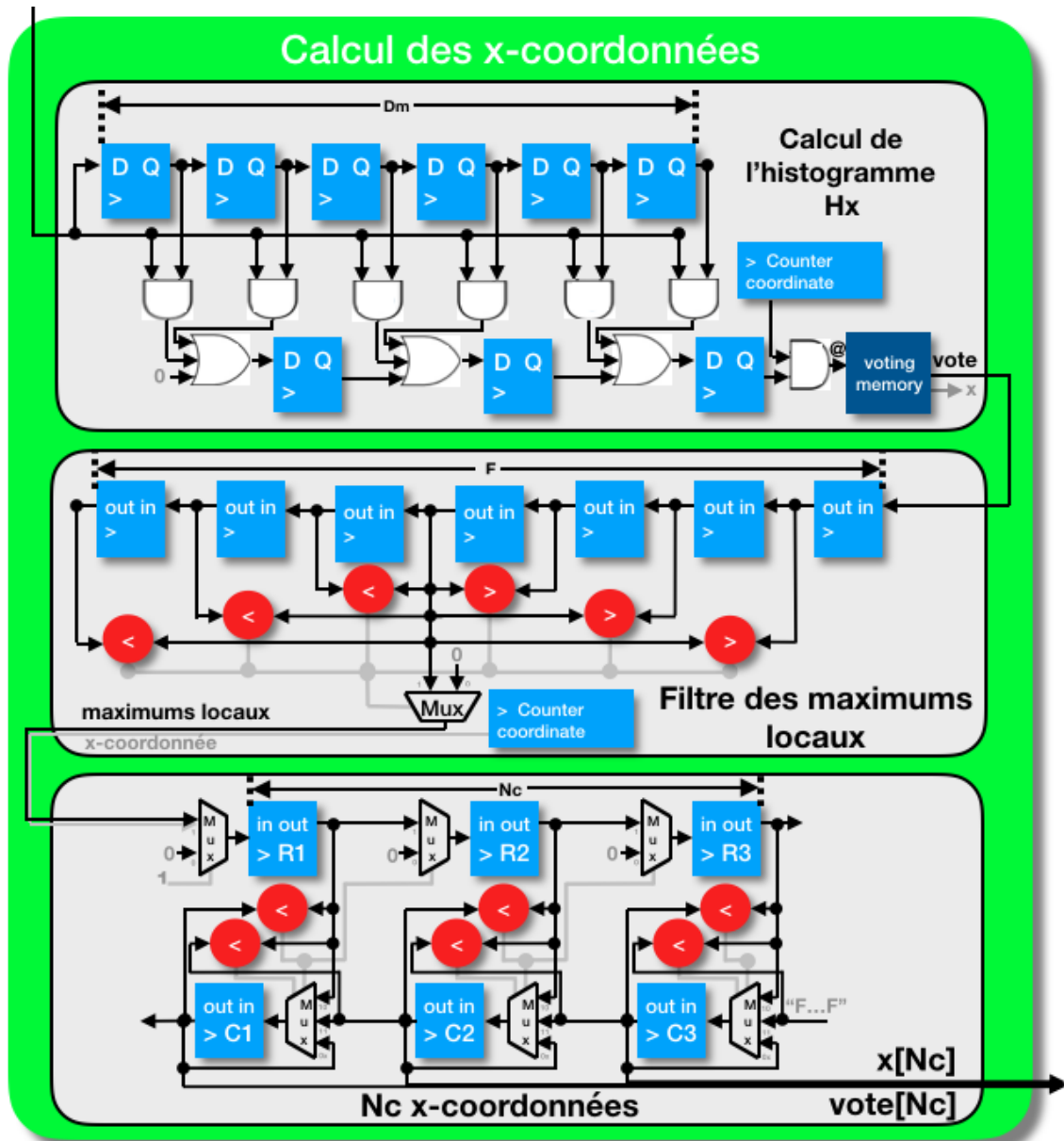


FIGURE 8.2 : Architecture optimisée en ressources pour le calcul des coordonnées en x.

Cette architecture est complètement pipelinée et composée de trois étages comme nous pouvons le voir dans la figure 8.2. Elle réalise le calcul de N_c coordonnées en x pour un diamètre maximal de D_m , avec $D_m > 1$.

Dans le premier étage, pour une image de taille $W * H$, l'image est lue pixel par pixel et par ligne. Chaque pixel est décalé en utilisant une bascule Flip-Flop d'un bit.

Un pixel d'entrée est multiplié directement avec les D_m pixels précédents de la même ligne. Les votes sont décalés avec les bascules Flip-Flop d'un bit et sont accumulés dans une seule mémoire.

Nous obtenons le même histogramme H_x que l'architecture proposée dans la section 7.2.2, un exemple de cette similarité est visible sur la figure 8.3.

Pour mémoriser les votes accumulés nous utilisons la même architecture que celle proposée dans la section 7.2.2 et visible dans la figure 7.8. Dans ce cas, les votes sont codés sur 20 bits et chaque mémoire est mise en œuvre dans une BRAM de 36 Kbits qui permet une lecture et une écriture simultanée en un seul cycle et en évitant les collisions.

Pour le deuxième et troisième étage nous utilisons la même architecture que la version 1DHT visible dans la figure 8.2 et décrit dans la section 7.2.1.

Les N_c coordonnées en y sont calculées avec une architecture similaire.

Au total, pour une image de taille $W * H$, $\frac{20(W+H)}{36K}$ BRAM de 36Kbits sont nécessaires pour le calcul des coordonnées en x et en y. Le calcul s'effectue en $W * H + W + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$ cycles, voir le tableau 8.2. Aucun module DSP n'est nécessaire.

Toutes les N_c coordonnées en x et en y sont combinées pour former N_c^2 centres candidats.

TABLE 8.2 : Ressources nécessaires pour le calcul des coordonnées en x et en y pour une image de taille $H * W$.

Module	Ressources nécessaires		Latence [cycles]
	BRAM de 36 Kbits	Bloc DSP	
Calcul des coordonnées en x et en y	$\frac{20(W+H)}{36K}$	0	$W * H + W + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$

Nous avons fait une comparaison des ressources nécessaires pour une image [FHD](#) à partir d'une estimation de cette architecture par rapport à la précédente, les résultats sont visibles dans le tableau 8.3. À partir de ces résultats nous remarquons que cette nouvelle architecture utilise 83% de BRAM en moins et aucun DSP, soit un gain de 100%, par rapport l'architecture précédente.

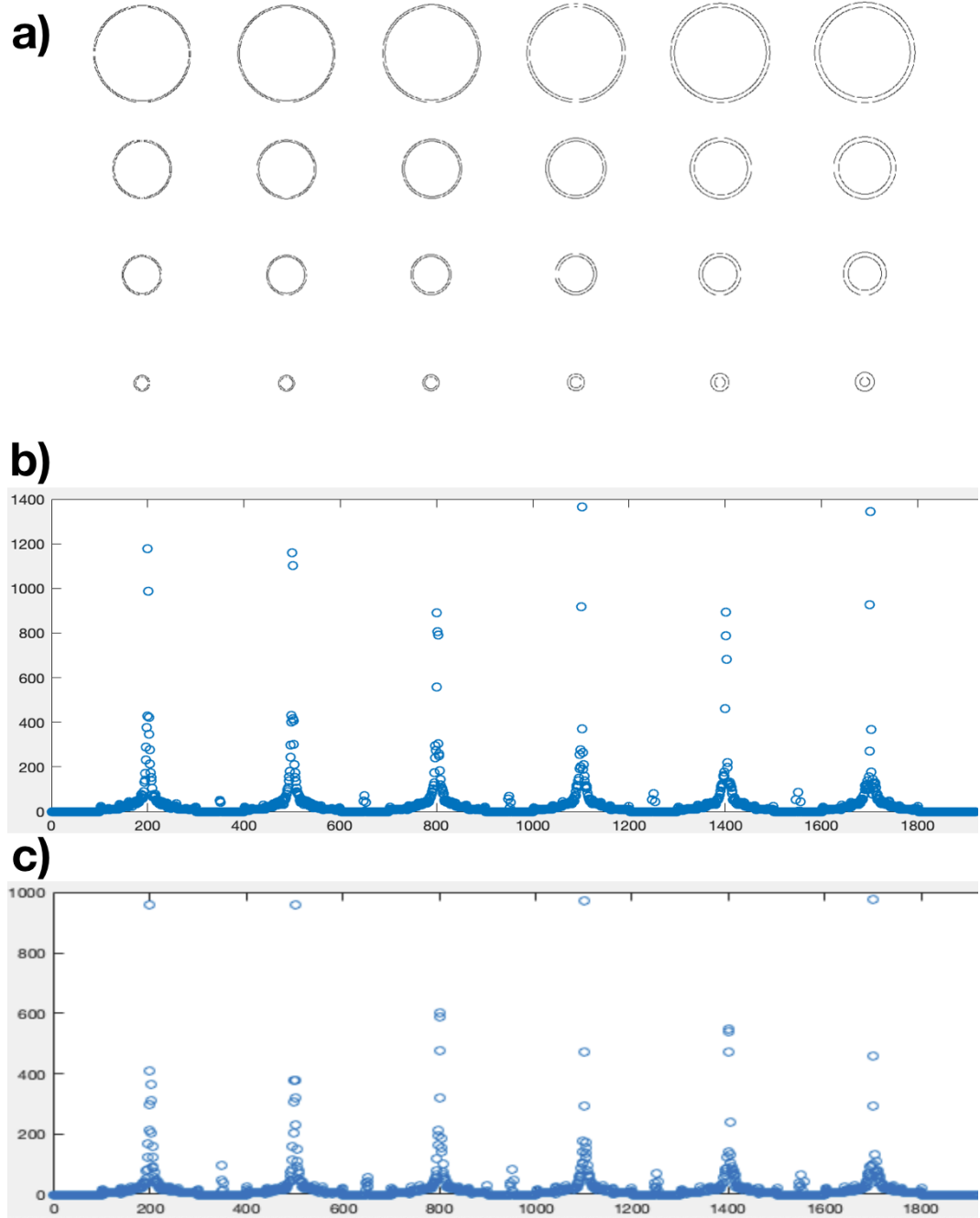


FIGURE 8.3 : a) Image de contour, histogramme H_x calculé par b) l'architecture x-coordonnées IP 1DHT et par c) la nouvelle architecture IP 1DHT optimisée. Dans les histogrammes l'axe x et y correspondent aux coordonnées en x et en y respectivement

TABLE 8.3 : Estimation des ressources nécessaires pour le calcul des coordonnées en x et en y.

Ressource	Architecture précédente	Nouvelle architecture	Réduction %
BRAM36Kb	12	2	83%
DSP	108	0	100%

8.3 Intégration de l'IP 1DHT-optimisée sur FPGA

Afin de déterminer si la nouvelle IP 1DHT-optimisée peut être intégrée dans une [CE](#), nous avons réalisé son implantation dans le circuit Spartan 7 XC7S15-CPGA196. Nous avons comparé les ressources utilisées pour notre architecture 1DHT-optimisée par rapport la version précédente. Les résultats de la mise en œuvre sont présentés dans le tableau 8.6.

TABLE 8.4 : Résultats d'implantation de l'IP 1DHT-optimisée sur le FPGA Xilinx Spartan 7 XC7S15.

Ressource	Disponibles	Utilisées par 1DHT-optimisée	Utilisées par 1DHT	Réduction
LUT	8000	7735(97%)	20637(258%)	63%
BRAM36K	10	10(100%)	78.5(785%)	87%
DSP	20	20(100%)	153(765%)	87%

À partir des résultats des tableaux 8.5 et 8.6, nous constatons que notre architecture peut être embarquée dans un [FPGA](#) de $8 * 8mm^2$ ce qui est compatible avec une intégration dans une [CE](#).

8.4 Intégration de notre IP 1DHT-optimisée dans une système sur puce

Pour valider notre architecture, nous l'avons mise en œuvre à l'aide de la carte de développement Digilent ZedBoard Zynq-7000 [ARM](#) / [FPGA](#) XC7Z020, afin de mesurer le temps d'exécution et de valider la fonctionnalité de notre architecture.

Pour la mise en œuvre de l'IP 1DHT-optimisée dans le Zynq nous exploitons le même pipeline au niveau système sur trois niveaux présenté dans la section 7.4.

Pour le calcul des centres candidats il faut $W * H + W + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$ cycles et pour le calcul des rayons il faut $W * H + 18 + \frac{D_m}{2}$ cycles. La latence total de la nouvelle IP AXI 1DHT-optimisée est égale à deux fois le temps de calcul des centres candidats.

Enfin, l'IP AXI 1DHT-optimisée peut détecter N_c^2 cercles candidats avec un diamètre maximum de D_m et séparés d'au minimum F pixels dans une image de

taille $W * H$. Ce calcul prends $2(W * H + W + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c)$ cycles de latence et $W * H + W + 2 + \frac{D_m}{2} + \frac{F}{2} + N_c$ cycles d'intervalle entre deux images.

La latence et l'intervalle sont directement liés à la taille de l'image.

Pour valider la fonctionnalité de notre architecture nous avons réalisé une mise en œuvre pour détecter $N_c^2 = 16$ cercles avec un diamètre maximum de $D_m = 108$ pixels séparés d'une distance minimale de $F = 40$ pixels dans des images [FHD](#). Les cercles détectés sont visibles dans la figure 8.4 et les détails de la mise en œuvre sont listés dans la table 8.5.

TABLE 8.5 : Détails de la mise en œuvre de l'IP "AXI 1DHT-optimisée".

Architecture	1DHT IP	AXI CHT IP IP core	Full Zynq System
LUT	7561	8087	8355
LUTRAM	176	179	283
FF	6951	7604	8651
BRAM (36Kb)	10	10	10
DSP	32	32	32
Freq. maximum[MHz]	140,10	135,46	135,46
Puiss consommée [mW]	68	76	1623
Freq. de test [MHz]	125		
Latence[ms]	33.24		
Intervalle[ms]	16.62		
Cadence[Mpixel/s]	125		
fps[frame/s]	60		

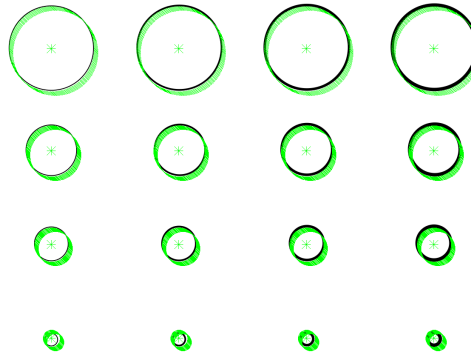


FIGURE 8.4 : Cercles détectés

Les performances obtenues sont listées dans le tableau 8.5. Notre IP peut fonctionner jusqu'à une fréquence de 140 MHz seule et jusqu'à une fréquence de 135 MHz intégrée dans un système. Nous avons validé notre architecture à une fréquence de 125 MHz pour des images [FHD](#). Cela nous donne une performance sur le traitement de chaque image de 16,62 ms.

Notre architecture peut traiter une image de taille 1920x1080 pixels en moins de

40 ms avec une cadence de 60 images per seconde. Mais qu'en est-il de sa consommation ?

8.5 Estimation de la consommation de l'IP optimisée

8.5.1 Avec un FPGA Spartan 7 XC7S15

Nous avons réalisé une estimation de la consommation énergétique de l'IP "AXI 1DHT-optimisée". Pour cela nous avons utilisé l'outil d'estimation de Puissance de Xilinx et avons obtenu une puissance estimée de 68mW. Pour estimer l'autonomie énergétique de cette architecture dans la capsule, nous avons réalisé une analyse énergétique.

Pour cette analyse nous avons considéré l'utilisation de deux batteries Energizer EPX76 qui peuvent être intégrées dans une [CE](#). Cette batterie a une tension nominale de 1,55 Volts et une capacité de 188mAh. Si les deux batteries sont en parallèle, la capacité totale est de $2 * 188mAh = 376mAh$ pour une tension de 1,55V. Notre architecture a été mise en œuvre sur le [FPGA](#) Xilinx Spartan 7 XC7S15 qui a une tension nominale de 0,95-1,0 volts. L'analyse nous permet de constater que nous obtenons une consommation énergétique de $\frac{68mW}{1V} = 68mA$. Avec la batterie EPX76 l'autonomie espérée est donc de $\frac{376mAh}{68mA} = 5,5h$. L'autonomie énergétique de notre architecture pour une cadence de 60 images par seconde est de 5,5 heures. Nous avons extrapolé l'autonomie pour différentes cadences, les autonomies sont visibles dans le tableau 8.6.

TABLE 8.6 : Autonomie énergétique estimée de l'IP "AXI 1DHT-optimisée" dans une [CE](#)

Cadence[fps]	4	16	25	32	60
1 image chaque [ms]	250	62,5	40	31,3	16,7
Autonomie [h]	82,5	20,6	13,2	10,3	5,5

À partir des résultats du tableau 8.6, avec une cadence de 25 images par seconde il est possible d'envisager une autonomie de 13,2 heures. Ce résultat va dans le sens de notre objectif d'avoir une autonomie supérieure à 10 heures.

Mais l'autonomie et la surface que nous avons obtenues correspondent à seulement une partie de notre chaîne, le calcul de la transformée de Hough. Cela ne prends pas en compte la consommation et la surface de l'ensemble de la chaîne de traitement et la consommation liée à l'acquisition de l'image et à la transmission des résultats.

Bien que la transformée de Hough corresponde à une des parties les plus critiques et coûteuses de notre chaîne de traitement, la surface du FPGA est déjà occupée à 99% et la consommation de la batterie est déjà proche de celle attendue de 10 heures. Cette analyse montre que pour l'intégration de notre chaîne complète, un [FPGA](#) n'est pas suffisant. Ceci nous oblige à développer un [ASIC](#).

8.5.2 Avec un ASIC en technologie 65 nm ST

Nous avons étudié la mise en œuvre de notre architecture IP 1DHT-optimisée sur un [ASIC](#) avec l'objectif d'estimer leur surface et leur consommation énergétique. Nous avons utilisé l'outil de développement RTL Compiler de Cadence avec une technologie CMOS de 65nm pour cela. Nous avons réglé la fréquence à 125MHz. Les résultats sont visibles dans le tableau 8.7.

TABLE 8.7 : Mise en œuvre de notre IP 1DHT-optimisée dans une technologie CMOS [ASIC](#) de 65 nm. Avec une fréquence de 125 MHz.

Architecture	Cellules	Surface(Cell+Physical+Net) [mm ²]	Consommation[mW]
1DHT-optimisée	165132	1,84	17,2
Calcul des centres candidats	97830	1,11	6,9
Calcul des rayons	67302	0,73	9,3

À partir des résultats du tableau 8.7, nous constatons que notre architecture avec une technologie CMOS 65 nm ST et pour une fréquence de 125MHz prend une surface de moins de 2 mm². La surface rectangulaire transversale de la capsule est de $8 \times 8 \text{ mm}^2 = 64 \text{ mm}^2$. Notre architecture pour le calcul de la transformée de Hough occupe moins de 4% de la surface disponible. De plus, la consommation énergétique est de 17 mW. Pour une tension nominale de 1,0 Volt nous obtenons une consommation énergétique de $\frac{17 \text{ mW}}{1 \text{ V}} = 17 \text{ mA}$. Avec la batterie EPX76 l'autonomie attendue est de $\frac{376 \text{ mAh}}{17 \text{ mA}} = 21,1 \text{ h}$. L'autonomie énergétique de notre architecture pour une cadence de 60 images per seconde est de 21,1 heures. Nous avons extrapolé l'autonomie pour une cadence de 25 images par seconde et nous obtenons une autonomie attendue de 50,6 heures. Cette autonomie ouvre la voie à la satisfaction de notre objectif d'avoir une autonomie supérieure à 10 heures en intégrant l'ensemble de la chaîne de traitement dans une [CE](#). Cette mise en œuvre est un premier modèle où on fait qu'une estimation globale dans lequel le taux d'activité n'a pas été mesuré et l'information de nombre des cellules SRAM n'as pas été extrait.

Synthèse et discussion

Dans ce chapitre nous avons analysé l’embarquabilité d’une partie de notre chaîne de traitement dans un système sur puce pour son intégration dans une CE.

Nous avons tout d’abord réalisé le profilage d’une version purement logicielle de notre chaîne de traitement pour estimer son temps d’exécution. Ceci nous a permis d’identifier les parties les plus lentes. Cette analyse a été réalisée sur un processeur embarqué ARM Cortex A9 avec une fréquence de 667 MHz en utilisant la librairie OpenCV. À l’aide de ce profilage nous avons obtenu le temps de traitement maximal de notre chaîne pour une image FHD qui est 3,0739s pour un nombre maximal de 13 ROI par image. Ce temps est 77 fois plus grand que notre objectif de traiter une image en 40ms. À partir de ce profilage nous avons identifié les deux parties le plus critiques de notre chaîne : le temps de calcul de la transformée de Hough dans l’étape d’extraction des ROI et le temps de calcul de descripteurs de Haralick dans l’étape de description des ROI. Nous avons décidé d’optimiser le temps le plus critique, c’est à dire le calcul de la transformée de Hough.

Avant toute réalisation d’une nouvelle architecture pour accélérer la transformée de Hough, nous avons réalisé un état de l’art de la réalisation d’architectures matérielles pour sa mise en œuvre. Ceci nous a permis de remarquer que les travaux existant avaient une utilisation importante des blocs de mémoires internes ou Blocs RAM (BRAM) et que principalement tous accélèrent le processus de vote dans la transformée de Hough. Ce processus est la partie plus lente de l’algorithme où il est accumulé des points d’intersection dans l’espace des paramètres de Hough qui correspond au nombre des cercles possibles. Le vote est l’action de trouver le maximum local qui sera considéré comme un cercle réel.

Pour mettre en œuvre ce processus deux différentes approches existent :

1. une première approche classique qui utilise l’algorithme original de la transformée de Hough où des équations paramétriques du cercle sont utilisées pour trouver le centre et le rayon des cercles [41], [38], [83], [52], [59]
2. une deuxième approche utilise une version modifiée de la transformée de Hough appelée transformée de Hough Unidimensionnelle (1DHT) [100].

Nous avons choisi d’étudier l’algorithme de 1DHT proposé par Goneid et al [45]

en 1997 et nous l'avons modifié pour optimiser les accès mémoires et la lecture des lignes et des colonnes de l'image dans le calcul des histogramme H_x et H_y et dans l'accélération du calcul des centres candidats.

Nous avons ensuite proposé une architecture pour la mise en œuvre de cet algorithme fonctionnant jusqu'à une fréquence de 149,16 MHz et traitant des images de taille 1920x1080 pixels en temps réel, c'est-à-dire en moins de 40 ms avec une cadence de 62 images par seconde. Notre architecture offre une vitesse 4 fois plus grande que l'architecture la plus rapide de l'état de l'art.

Afin de déterminer si notre architecture peut être intégrée dans une [CE](#), nous avons choisi comme cible d'intégration un FPGA Xilinx Spartan 7 XC7S15-CPGA196, qui possède une taille de $8 * 8mm^2$ proche de celle nécessaire pour une [CE](#) et avons remarqué qu'elle ne peut pas être intégrée dans un cette cible. Une analyse nous a indiqué que les BRAM et les DSP sont les ressources les plus utilisées. Nous avons ré-optimisé notre architecture pour permettre son intégration dans ce FPGA et dans une [CE](#).

Nous avons modifié l'architecture pour les calcul des coordonnées en x et en y pour limiter l'utilisation des BRAM et des blocs DSP. Cette nouvelle architecture peut supporter jusqu'à une fréquence de 140 MHz seule et de 135 MHz dans le système global, elle traite des images [FHD](#) en 16,62 ms.

Notre architecture peut traiter une image de taille 1920x1080 pixels en moins de 40 ms avec une cadence de 60 images par seconde et peut être intégrée dans une [CE](#). Une estimation de la puissance a été faite, nous avons obtenu une puissance estimée de 68 mWatt. À partir de cette puissance, nous avons estimé l'autonomie énergétique. Cette estimation a indiqué une autonomie de 13,2 heures pour une cadence de 25 images par seconde. Cette autonomie satisfait l'objectif d'avoir une autonomie supérieure à 10 heures dans le cadre unique du calcul de la transformée de Hough qui n'est qu'un traitement de notre chaîne.

Pour intégrer notre chaîne dans une [CE](#) nous avons décidé de développer un [ASIC](#) fonctionnant à une fréquence de 125 MHz avec une technologie CMOS 65nm. Nous avons obtenu une surface de moins de $2 mm^2$ et une autonomie de 50,6h pour une cadence de 25 images par seconde pour la transformée de Hough, ce qui ouvre la voie d'une intégration complète de notre chaîne sur un [ASIC](#) en satisfaisant les contraintes liées à une [CE](#).

Quatrième partie

CONCLUSION

Chapitre 9

Contributions et perspectives

One, remember to look up at the stars and not down at your feet. Two, never give up work. Work gives you meaning and purpose and life is empty without it. Three, if you are lucky enough to find love, remember it is there and don't throw it away.

Stephen Hawking

9.1 Contributions

Les travaux présentés dans ce manuscrit de thèse ont abordé les limites des CE, c'est-à-dire leurs incapacité à réaliser un diagnostic précoce. La solution proposée pour dépasser cette limite a engendré d'autres problématiques scientifiques liées à l'intégration et l'embarquabilité d'algorithmes de traitement d'images au sein d'une CE.

En partant de l'étude des limitations des méthodes de dépistage et de diagnostic actuelles, nous avons défini des caractéristiques souhaitables d'un dispositif médical pour le diagnostic précoce du CCR en introduisant un concept nouveau et original : celui d'une CE intelligente capable de sélectionner des images pouvant contenir des polypes, une iCE.

Nous avons proposé des contributions pour réaliser cette iCE afin de développer un SVSoC intégré en son sein. Pour cela, dans une première partie une chaîne de traitement, s'inspirant des travaux décrits dans [78], pour la reconnaissance automatique des polypes a été proposée et validée avec une base de données publique. Ensuite dans un second temps, la chaîne de traitement a été mise en œuvre dans un processeur embarqué pour évaluer chaque partie de la chaîne. La partie la plus chronophage de la chaîne de traitement a été identifiée, c'est la transformée de Hough.

Nous avons proposé une architecture matérielle optimisée pour la transformée de Hough. Cette architecture peut traiter des images HD en temps réel et peut être intégrée dans un FPGA de $8 * 8mm^2$, ce qui correspond à la surface d'une CE, tout en possédant une autonomie de plus de 10 heures.

A partir de ces contributions nous pouvons conclure que :

- Les descripteurs des texture et de luminosité sont des descripteurs pertinents et discriminants pour la reconnaissance des polypes sur des ROI même quand la ROI ne contient qu'une partie d'un polype ;
- Les arbres flous ont une stabilité et une capacité élevées pour la reconnaissance des polypes ;
- La forêt floue permet d'améliorer les performances des arbres flous ;
- En prenant en compte l'information spatio-temporelle d'une ROI contenant un polype, il est possible d'améliorer les performances de détection ;
- L'entraînement des forêts floues pour reconnaître des ROI contenant au moins une partie d'un polype permet d'augmenter les performances de détection ;
- Les parties les plus critiques de notre chaîne de traitement sont le calcul de la transformée de Hough et l'extraction des descripteurs de texture.

- La transformée de Hough est intégrable dans une [CE](#) pour le traitement des images [FHD](#) en temps réel avec une autonomie énergétique de plus de 10 heures.
- Pour intégrer la chaîne complète de traitement, il est nécessaire de développer un [ASIC](#).

9.2 Perspectives

Une des difficultés rencontrées lors de nos travaux est liée au peu de disponibilité de base de données annotées dans le domaine médical et notamment dans celui lié au [CCR](#). Une des perspectives que nous avons identifiée est celle concernant la création d'une base de données d'images de capsule contenant des polypes de plusieurs types, incluant les déplacements de la caméra avec une annotation qui indique la présence ou l'absence de polypes, mais aussi le type de polype et sa localisation dans le côlon ou le rectum.

Une autre de nos perspectives est l'étude du suivi des polypes sur des images de capsule endoscopique. Cette étude nous permettra d'évaluer différentes méthodes d'estimation de mouvement pour une meilleure intégration dans une [CE](#).

Une autre perspective est le développement d'une méthode pour l'agrégation des résultats de plusieurs images afin de prendre une décision sur l'existence d'un polype sur une séquence d'images et non plus une seule image.

Dans le profilage réalisé, nous avons aussi identifié que le calcul des descripteurs de Haralick était aussi un calcul chronophage, une future étape est de réaliser l'état de l'art pour l'accélération de ce calcul et de proposer, si cela est pertinent, une architecture permettant de calculer ces descripteurs tout en respectant les contraintes liées à la [iCE](#).

Dans nos travaux nous avons mise en œuvre la partie la plus critique de notre chaîne de traitement sur [FPGA](#) et sur [ASIC](#). Une perspective est la conception d'une architecture matérielle pour l'intégration de la totalité de la chaîne de traitement est aussi le développement d'un [ASIC](#) qui intègre cette architecture pour réaliser un premier prototype d'une [iCE](#). Ce prototype nous permettra de réaliser des tests et des mesures in-vivo pour obtenir une validation clinique.

Du point de vue scientifique et technologique nous avons validé un dispositif médical pour la détection automatique des polypes. Il est nécessaire de passer maintenant vers des études cliniques pour évaluer son impact dans le diagnostic précoce du [CCR](#).

Cinquième partie

ANNEXES

Annexe A

Annexes

A.1 Descripteurs extraits à partir de la matrice de co-occurrence et de l'histogramme de luminosité

À partir de la matrice de Co-occurrence $M_{256 \times 256}$ et l'histogramme de luminosité normalisée $H_{(256)}$ 26 descripteurs [49],[3] de texture et luminosité ont été extraits. Les descripteurs sont les suivants :

A.1.1 Descripteurs calculés à partir de l'histogramme de luminosité H :

1. mean

$$f_1 = \mu = \sum_{i=0}^{255} iH(i) \quad (\text{A.1})$$

2. Variance

$$f_2 = \sigma^2 = \sum_{i=0}^{255} (i - \mu)^2 H(i) \quad (\text{A.2})$$

3. Skewness

$$f_3 = \sigma^{-3} \sum_{i=0}^{255} (i - \mu)^3 H(i) \quad (\text{A.3})$$

4. kurtosis

$$f_4 = \sigma^{-4} \sum_{i=0}^{255} (i - \mu)^4 H(i) \quad (\text{A.4})$$

A.1.2 Descripteurs calculés à partir de la Matrice de Co-occurrence M :

5. Autocorrélation

$$f_5 = \sum_{i=0}^{255} \sum_{j=0}^{255} ijM(i, j) \quad (\text{A.5})$$

6. Contraste

$$f_6 = \sum_{i=0}^{255} \sum_{j=0}^{255} (i - j)^2 M(i, j) \quad (\text{A.6})$$

7. Corrélacion

$$f_7 = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{ijM(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (\text{A.7})$$

Où :

$$\mu_x = \sum_{i=0}^{255} \sum_{j=0}^{255} iM(i, j) \quad (\text{A.8})$$

$$\mu_y = \sum_{i=0}^{255} \sum_{j=0}^{255} jM(i, j) \quad (\text{A.9})$$

$$\sigma_x = \sum_{i=0}^{255} \sum_{j=0}^{255} (i - \mu_x)^2 M(i, j) \quad (\text{A.10})$$

$$\sigma_y = \sum_{i=0}^{255} \sum_{j=0}^{255} (j - \mu_y)^2 M(i, j) \quad (\text{A.11})$$

8. Corrélacion proposée par Chiolerio et al.[32]

$$f_8 = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{(i - \mu_x)(j - \mu_y)M(i, j)}{\sigma_x \sigma_y} \quad (\text{A.12})$$

9. Dissimilarité

$$f_9 = \sum_{i=0}^{255} \sum_{j=0}^{255} |i - j| M(i, j) \quad (\text{A.13})$$

10. Cluster Shade

$$f_{10} = \sum_{i=0}^{255} \sum_{j=0}^{255} (i + j - \mu_x - \mu_y)^3 M(i, j) \quad (\text{A.14})$$

11. Cluster Prominence

$$f_{11} = \sum_{i=0}^{255} \sum_{j=0}^{255} (i + j - \mu_x - \mu_y)^4 M(i, j) \quad (\text{A.15})$$

12. Energie Matlab

$$f_{12} = \sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j)^2 \quad (\text{A.16})$$

13. Entropie

$$f_{13} = - \sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j) \log_2 M(i, j) \quad (\text{A.17})$$

14. Probabilité Maximale

$$f_{14} = \text{Max}_{i,j}(M(i, j)) \quad (\text{A.18})$$

15. Homogénéité

$$f_{15} = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{M(i, j)}{1 + |i - j|} \quad (\text{A.19})$$

16. IDM (Inverse Difference Moment) [10]

$$f_{16} = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{M(i, j)}{1 + (i - j)^2} \quad (\text{A.20})$$

17. Somme des racines carrés : variance

$$f_{17} = \sum_{i=0}^{2x255} (i - \mu)^2 M(i, j) \quad (\text{A.21})$$

18. Somme des moyenne

$$f_{18} = \sum_{i=0}^{2x255} i M_{x+y}(i) \quad (\text{A.22})$$

Où :

$$M_{x+y}(k) = \frac{\sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j)}{|i + j| = k} \quad k = 0, 2, 3, \dots, 2x255 \quad (\text{A.23})$$

19. Somme des variances

$$f_{19} = \sum_{i=0}^{2x255} (i - f_{18})^2 M_{x+y}(i) \quad (\text{A.24})$$

20. Somme des entropies

$$f_{20} = \sum_{i=0}^{255} M_{x+y} \log M_{x+y} \quad (\text{A.25})$$

21. Différence des variances

$$f_{21} = \sum_{i=0}^{255} (i - \mu_{M_{x-y}})^2 M_{x-y}(i) \quad (\text{A.26})$$

Où :

$$M_{x-y}(k) = \frac{\sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j)}{|i - j| = k} \quad k = 0, 2, 3, \dots, 255 - 1 \quad (\text{A.27})$$

$$\mu_{M_{x-y}} = \sum_{i=0}^{255} i M_{x-y}(i) \quad (\text{A.28})$$

22. Différence des entropies

$$f_{22} = \sum_{i=0}^{255} M_{x-y} \log M_{x-y} \quad (\text{A.29})$$

23-24. Information mesure de corrélation

$$f_{23} = \frac{MXY - MXY1}{\max(MX, MY)} \quad (\text{A.30})$$

$$f_{24} = (1 - \exp(-2(MXY2 - MXY)))^{\frac{1}{2}} \quad (\text{A.31})$$

Où :

$$M_x(i) = \sum_{j=0}^{255} M(i, j) \quad (\text{A.32})$$

$$M_y(i) = \sum_{j=0}^{255} M(i, j) \quad (\text{A.33})$$

$$MX = - \sum_{i=0}^{255} M_x(i) \log M_x(i) \quad (\text{A.34})$$

$$MY = - \sum_{i=0}^{255} M_y(i) \log M_y(j) \quad (\text{A.35})$$

$$MXY = - \sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j) \log M(i, j) \quad (\text{A.36})$$

$$MXY1 = - \sum_{i=0}^{255} \sum_{j=0}^{255} M(i, j) \log M_x(i) M_y(j) \quad (\text{A.37})$$

$$MXY2 = - \sum_{i=0}^{255} \sum_{j=0}^{255} M_x(i) M_y(j) \log M_x(i) M_y(j) \quad (\text{A.38})$$

25. Inverse différence normalisée

$$f_{25} = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{M(i, j)}{1 + (i - j)^2} \quad (\text{A.39})$$

26. Inverse différence moment

$$f_{26} = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{M(i, j)}{1 + (i - j)^2} \quad (\text{A.40})$$

Bibliographie

- [1] CapsoCam Plus® | CapsoVision, <https://www.capsovision.com/products/capsocam-plus>.
- [2] ENDOCAPSULE 10 | Olympus America, <https://medical.olympusamerica.com/products/endocapsule>.
- [3] GLCM texture features. <https://fr.mathworks.com/matlabcentral/fileexchange/22187>.
- [4] A High-Performance Reconfigurable Accelerator for Convolutional Neural Networks | Proceedings of the 3rd International Conference on Multimedia Systems and Signal Processing.
- [5] JINSHAN Science & Technology Capsule Endoscopy, <http://english.jinshangroup.com/capsuleendoscopy.html>.
- [6] La Sociedad Americana Contra El Cancer | Información sobre cáncer de seno, colon, pulmón, piel, y otros. <https://www.cancer.org/es>.
- [7] PillCam™ COLON 2 System | Medtronic, <https://www.medtronic.com/covidien/en-us/products/capsule-endoscopy/>.
- [8] Wireless Capsule Endoscopy, [https://www.giejournal.org/article/s0016-5107\(13\)02091-9/abstract](https://www.giejournal.org/article/s0016-5107(13)02091-9/abstract).
- [9] Zedboard, <http://www.zedboard.org>.
- [10] AXI Reference Guide, https://www.xilinx.com/support/documentation/ip_documentation/ug761_axi_reference_guide.pdf, 1973.
- [11] AXI Reference Guide, https://www.xilinx.com/support/documentation/ip_documentation/ug761_axi_reference_guide.pdf, 2011.
- [12] Cancer Today, <http://gco.iarc.fr/today/home>, 2018.
- [13] Zynq-7000 SoC Data Sheet : Overview (DS190), https://www.xilinx.com/support/documentation/data_sheets/ds190-zynq-7000-overview.pdf, 2018.
- [14] AARONSON, N. K., AHMEDZAI, S., BERGMAN, B., BULLINGER, M., CULL, A., DUEZ, N. J., FILIBERTI, A., FLECHTNER, H., FLEISHMAN, S. B., DE HAES, J. C. J. M., KAASA, S., KLEE, M., OSOBA, D., RAZAVI, D., ROFE, P. B., SCHRAUB, S., SNEEUW, K., SULLIVAN, M., AND TAKEDA, F. The

- European Organization for Research and Treatment of Cancer QLQ-C30 : A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *J Natl Cancer Inst* 85, 5 (Mar. 1993), 365–376.
- [15] ACHARYA, T., AND RAY, A. K. *Image Processing : Principles and Applications*. John Wiley & Sons, Oct. 2005.
- [16] ADLER, S. N., AND METZGER, Y. C. PillCam COLON Capsule Endoscopy : Recent Advances and New Insights. *Ther. Adv Gastroenterol* 4, 4 (July 2011), 265–268.
- [17] AKOUSHIDEH, A. R., SHAHBAHRAMI, A., AND MAYBODI, B. M.-N. High Performance Implementation of Texture Features Extraction Algorithms Using FPGA Architecture. *J Real-Time Image Proc* 9, 1 (Mar. 2014), 141–157.
- [18] AND. Design and Implementation of the Tree-Based Fuzzy Logic Controller. *IEEE Trans Syst Man Cybern Part B Cybern* 27, 3 (June 1997), 475–487.
- [19] ANDRAKA, R. A survey of CORDIC algorithms for FPGA based computers. In *Proceedings of the 1998 ACM/SIGDA Sixth International Symposium on Field Programmable Gate Arrays - FPGA '98* (Monterey, California, United States, 1998), ACM Press, pp. 191–200.
- [20] ANGERMANN, Q., BERNAL, J., SANCHEZ-MONTES, C., HAMMAMI, M., FERNANDEZ-ESPARRACH, G., DRAY, X., ROMAIN, O., SANCHEZ, F. J., AND HISTACE, A. Towards Real-Time Polyp Detection in Colonoscopy Videos : Adapting Still Frame-Based Methodologies for Video Sequences Analysis. In *Computer Assisted and Robotic Endoscopy and Clinical Image-Based Procedures* (2017), M. J. Cardoso, T. Arbel, X. Luo, S. Wesarg, T. Reichl, M. A. Gonzalez Ballester, J. McLeod, K. Drechsler, T. Peters, M. Erdt, K. Mori, M. G. Linguraru, A. Uhl, C. Oyarzun Laura, and R. Shekhar, Eds., Lecture Notes in Computer Science, Springer International Publishing, pp. 29–41.
- [21] ANTALOVIC, D. Review of the Hough Transform Method, With an Implementation of the Fast Hough Variant for Line Detection. *DCS* (2008).
- [22] AXON, A., DIEBOLD, M. D., FUJINO, M., FUJITA, R., GENTA, R. M., GONVERS, J. J., GUELHUD, M., INOUE, H., JUNG, M., KASHIDA, H., KUDO, S., LAMBERT, R., LIGHTDALE, C., NAKAMURA, T., NEUHAUS, H., NIWA, H., OGOSHI, K., REY, J. F., RIDDELL, R., SASAKO, M., SHIMODA, T., SUZUKI, H., TYTGAT, G. N. J., WANG, K. K. N., WATANABE, H., YAMAKAWA, T., AND YOSHIDA, S. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Endoscopy*. 37, 6 (June 2005), 570–578.
- [23] BERNAL, J., SÁNCHEZ, F. J., FERNÁNDEZ-ESPARRACH, G., GIL, D., RODRÍGUEZ, C., AND VILARIÑO, F. WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy : Validation vs. Saliency Maps from Physicians. *Comput. Med. Imaging Graph.* 43 (July 2015), 99–111.

-
- [24] BERNAL, J., SÁNCHEZ, J., AND VILARINO, F. Towards Automatic Polyp Detection with a Polyp Appearance Model. *Pattern Recognit* 45, 9 (2012), 3166–3182.
- [25] BERNAL, J., TAJKBAKSH, N., SÁNCHEZ, F. J., MATUSZEWSKI, B. J., CHEN, H., YU, L., ANGERMANN, Q., ROMAIN, O., RUSTAD, B., BALASINGHAM, I., POGORELOV, K., CHOI, S., DEBARD, Q., MAIER-HEIN, L., SPEIDEL, S., STOYANOV, D., BRANDAO, P., CÓRDOVA, H., SÁNCHEZ-MONTES, C., GURUDU, S. R., FERNÁNDEZ-ESPARRACH, G., DRAY, X., LIANG, J., AND HISTACE, A. Comparative Validation of Polyp Detection Methods in Video Colonoscopy : Results From the MICCAI 2015 Endoscopic Vision Challenge. *IEEE Trans Med Imaging* 36, 6 (June 2017), 1231–1249.
- [26] BOSER, B. E., GUYON, I. M., AND VAPNIK, V. N. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (New York, NY, USA, 1992), COLT '92, ACM, pp. 144–152. event-place : Pittsburgh, Pennsylvania, USA.
- [27] BOURBAKIS, N., GIAKOS, G., AND KARARGYRIS, A. Design of New-Generation Robotic Capsules for Therapeutic and Diagnostic Endoscopy. In *2010 IEEE International Conference on Imaging Systems and Techniques* (July 2010), pp. 1–6.
- [28] BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A., AND JEMAL, A. Global Cancer Statistics 2018 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 68, 6 (2018), 394–424.
- [29] BRENNER, H., AND TAO, S. Superior Diagnostic Performance of Faecal Immunochemical Tests for Haemoglobin in a Head-to-Head Comparison with Guaiac Based Faecal Occult Blood Test among 2235 Participants of Screening Colonoscopy. *Eur. J. Cancer* 49, 14 (Sept. 2013), 3049–3054.
- [30] BUADES, A., LE, T., MOREL, J.-M., AND VESE, L. Cartoon+Texture Image Decomposition. *Image Process Line 1* (Sept. 2011), 200–207.
- [31] CANNY, J. A Computational Approach to Edge Detection. *IEEE Trans Pattern Anal Mach Intell PAMI-8*, 6 (Nov. 1986), 679–698.
- [32] CHIOLERIO, A., CHIODONI, A., AND ALLIA, P. Elemental distribution and morphological analysis of layered metallic systems : Application to co–sn evaporated multilayers. *Thin Solid Films* 516, 23 (2008), 8453–8461.
- [33] DANG, T. H. *Discrimination measures and theirs applications in inductif learning*. Theses, Université Pierre et Marie Curie - Paris VI, July 2007.
- [34] DAVID, E., BOIA, R., MALAESCU, A., AND CARNU, M. Automatic colon polyp detection in endoscopic capsule images. *International Symposium on Signals, Circuits and Systems ISSCS2013* (2013).

-
- [35] DE WIJKERSLOOTH, T. R., STOOP, E. M., BOSSUYT, P. M., MEIJER, G. A., VAN BALLEGOOIJEN, M., VAN ROON, A. H. C., STEGEMAN, I., KRAAIJENHAGEN, R. A., FOCKENS, P., VAN LEERDAM, M. E., DEKKER, E., AND KUIPERS, E. J. Immunochemical Fecal Occult Blood Testing Is Equally Sensitive for Proximal and Distal Advanced Neoplasia. *Am J Gastroenterol* 107, 10 (Oct. 2012), 1570–1578.
- [36] DIK, V. K., MOONS, L. M., AND SIERSEMA, P. D. Endoscopic Innovations to Increase the Adenoma Detection Rate during Colonoscopy. *World J Gastroenterol* 20, 9 (Mar. 2014), 2200–2211.
- [37] D’ORAZIO, L., BARTOLI, A., BAETZ, A., BEORCHIA, S., CALVARY, G., CHABANE, Y., CHADEBECQ, F., COLLINS, T., LAURILLAU, Y., MARTINS-BALTAR, L., MOHAMAD, B., PONCHON, T., REY, C., TILMANT, C., AND TORTI, S. Multimodal and Multimedia Image Analysis and Collaborative Networking for Digestive Endoscopy. *IRBM* 35, 2 (Apr. 2014), 88–93.
- [38] ELHOSSINI, A., AND MOUSSA, M. Memory Efficient Fpga Implementation of Hough Transform for Line and Circle Detection. In *Electrical & Computer Engineering (CCECE), 2012 25th IEEE Canadian Conference On* (2012), IEEE, pp. 1–5.
- [39] ELIAKIM, R., YASSIN, K., NIV, Y., METZGER, Y., LACHTER, J., GAL, E., SAPOZNIKOV, B., KONIKOFF, F., LEICHTMANN, G., FIREMAN, Z., KOPELMAN, Y., AND ADLER, S. N. Prospective Multicenter Performance Evaluation of the Second-Generation Colon Capsule Compared with Colonoscopy. *Endoscopy* 41, 12 (Dec. 2009), 1026–1031.
- [40] ELSAFI, S. H., ALQAHTANI, N. I., ZAKARY, N. Y., AND AL ZAHRANI, E. M. The Sensitivity, Specificity, Predictive Values, and Likelihood Ratios of Fecal Occult Blood Test for the Detection of Colorectal Cancer in Hospital Settings. *Clin Exp Gastroenterol* 8 (Sept. 2015), 279–284.
- [41] FERHAT-TALEB ALIM, F., MESSAOUDI, K., SEDDIKI, S., AND KERDJIDJ, O. Modified Circular Hough Transform Using FPGA. In *2012 24th International Conference on Microelectronics (ICM)* (Dec. 2012), IEEE, pp. 1–4.
- [42] FERHAT-TALEB ALIM, F., MESSAOUDI, K., SEDDIKI, S., AND KERDJIDJ, O. Modified Circular Hough Transform Using FPGA. In *2012 24th International Conference on Microelectronics (ICM)* (Dec. 2012), IEEE, pp. 1–4.
- [43] FILIP, D., YADID-PECHT, O., MUENCH, G., MINTCHEV, M. P., AND ANDREWS, C. N. Suture Marker Lesion Detection in the Colon by Self-Stabilizing and Unmodified Capsule Endoscopes : Pilot Study in Acute Canine Models. *Gastrointest. Endosc.* 77, 2 (Feb. 2013), 272–279.
- [44] FISHER, L. R., AND HASLER, W. L. New Vision in Video Capsule Endoscopy : Current Status and Future Directions. *Nat Rev Gastroenterol Hepatol* 9, 7 (July 2012), 392–405.

- [45] GONEID, A., EL-GINDI, S., AND SEWISY, A. A method for the Hough transform detection of circles and ellipses using a 1-dimensional array. In *Computational Cybernetics and Simulation 1997 IEEE International Conference on Systems, Man, and Cybernetics* (Oct. 1997), vol. 4, pp. 3154–3157 vol.4.
- [46] GOPI, V. P., AND PALANISAMY, P. Capsule Endoscopic Image Denoising Based on Double Density Dual Tree Complex Wavelet Transform. 14.
- [47] HÄFNER, M., LIEDLGRUBER, M., AND UHL, A. POCS-Based Super-Resolution for HD Endoscopy Video Frames. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems* (June 2013), pp. 185–190.
- [48] HARA, A. K., JOHNSON, C. D., REED, J. E., AHLQUIST, D. A., NELSON, H., MACCARTY, R. L., HARMSSEN, W. S., AND ILSTRUP, D. M. Detection of Colorectal Polyps with CT Colography : Initial Assessment of Sensitivity and Specificity. *Radiology* 205, 1 (Oct. 1997), 59–65.
- [49] HARALICK, R. M., SHANMUGAM, K., AND DINSTEN, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern. SMC-3*, 6 (Nov. 1973), 610–621.
- [50] HWANG, S. Bag Of Visual Words Approach Based on SURF Features to Polyp Detection in Wireless Capsule Endoscopy Videos. 4.
- [51] IAKOVIDIS, D. K., AND KOULAOUZIDIS, A. Software for Enhanced Video Capsule Endoscopy : Challenges for Essential Progress. *Nat Rev Gastroenterol Hepatol* 12, 3 (2015), 172–186.
- [52] IRWANSYAH, A., IBRAHEEM, O. W., HAGEMEYER, J., PORRMANN, M., AND RUECKERT, U. FPGA-Based Circular Hough Transform with Graph Clustering for Vision-Based Multi-Robot Tracking. In *2015 International Conference on ReConFigurable Computing and FPGAs (ReConFig)* (Dec. 2015), IEEE, pp. 1–8.
- [53] KARARGYRIS, A., AND BOURBAKIS, N. Identification of polyps in Wireless Capsule Endoscopy videos using Log Gabor filters. In *2009 IEEE/NIH Life Science Systems and Applications Workshop* (Apr. 2009), pp. 143–147.
- [54] KARARGYRIS, A., AND BOURBAKIS, N. Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos. *IEEE Trans Biomed Eng* 58, 10 (Oct. 2011), 2777–2786.
- [55] KODOGIANNIS, V., AND BOULOUGOURA, M. An Adaptive Neurofuzzy Approach for the Diagnosis in Wireless Capsule Endoscopy Imaging. *Int J Inf Technol* 13 (2007), 46–56.
- [56] KRIZHEVSKY, A., SUTSKEVER, I., AND HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

-
- [57] KULPA, Z. On the properties of discrete circles, rings, and disks. *Computer Graphics and Image Processing* 10, 4 (Aug. 1979), 348–365.
 - [58] KUMAR, V., ASATI, A., AND GUPTA, A. Hardware Accelerators for Iris Localization. *J Sign Process Syst* 90, 4 (Apr. 2018), 655–671.
 - [59] KUMAR, V., ASATI, A., AND GUPTA, A. Hardware Accelerators for Iris Localization. *J Sign Process Syst* 90, 4 (Apr. 2018), 655–671.
 - [60] LANDAHL, H. D., MCCULLOCH, W. S., AND PITTS, W. A Statistical Consequence of the Logical Calculus of Nervous Nets. *Bull. Math. Biophys.* 5, 4 (Dec. 1943), 135–137.
 - [61] LANSDORP-VOGELAAR, I., VAN BALLEGOOIJEN, M., ZAUBER, A. G., HABBEMA, J. D. F., AND KUIPERS, E. J. Effect of Rising Chemotherapy Costs on the Cost Savings of Colorectal Cancer Screening. *J Natl Cancer Inst* 101, 20 (Oct. 2009), 1412–1422.
 - [62] LECUN, Y. Comparing Different Neural Network Architectures for Classifying Handwritten Digits. *IJCNN Proc N IEEE Hebb 1949* (1989).
 - [63] LEE, Y. Y., ERDOGAN, A., AND RAO, S. S. C. How to Assess Regional and Whole Gut Transit Time With Wireless Motility Capsule. *J Neurogastroenterol Motil* 20, 2 (Apr. 2014), 265–270.
 - [64] LEVIN, T. R., CORLEY, D. A., JENSEN, C. D., SCHOTTINGER, J. E., QUINN, V. P., ZAUBER, A. G., LEE, J. K., ZHAO, W. K., UDALTSOVA, N., GHAI, N. R., LEE, A. T., QUESENBERRY, C. P., FIREMAN, B. H., AND DOUBENI, C. A. Effects of Organized Colorectal Cancer Screening on Cancer Incidence and Mortality in a Large Community-Based Population. *Gastroenterology* 155, 5 (Nov. 2018), 1383–1391.e5.
 - [65] LEVY, B. T., BAY, C., XU, Y., DALY, J. M., BERGUS, G., DUNKELBERG, J., AND MOSS, C. Test Characteristics of Fecal Immunochemical Tests (FIT) Compared with Optical Colonoscopy Revised JMS-14-003.R2. *J Med Screen* 21, 3 (Sept. 2014), 133–143.
 - [66] LI, B., AND MENG, M. Q.-H. Automatic Polyp Detection for Wireless Capsule Endoscopy Images. *Expert Syst Appl* 39, 12 (2012), 10952–10958.
 - [67] MAMONOV, A. V., FIGUEIREDO, I. N., FIGUEIREDO, P. N., AND TSAI, Y.-H. R. Automated polyp detection in colon capsule endoscopy. *IEEE Trans. Med. Imaging* 33, 7 (2014), 1488–1502.
 - [68] MARSALA, C. *Apprentissage Inductif En Presence de Donnees Imprecises : Construction et Utilisation d’arbres de Decision Flous*. Thesis, Paris 6, Jan. 1998.
 - [69] MARSALA, C., AND BOUCHON-MEUNIER, B. Fuzzy partitioning using mathematical morphology in a learning scheme. In *Proceedings of IEEE 5th International Fuzzy Systems* (Sept. 1996), vol. 2, pp. 1512–1517 vol.2.

-
- [70] MORGAN, J. N., AND SONQUIST, J. A. Problems in the Analysis of Survey Data, and a Proposal. *J. Am. Stat. Assoc.* 58, 302 (June 1963), 415–434.
- [71] MUKHOPADHYAY, P., AND CHAUDHURI, B. B. A Survey of Hough Transform. *Pattern Recognit* 48, 3 (2015), 993–1010.
- [72] NAWARATHNA, R. D., OH, J., YUAN, X., LEE, J., AND TANG, S. J. Abnormal Image Detection Using Texton Method in Wireless Capsule Endoscopy Videos. In *Medical Biometrics* (2010), D. Zhang and M. Sonka, Eds., Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 153–162.
- [73] OKUHATA, H., NAKAMURA, H., HARA, S., TSUTSUI, H., AND ONOYE, T. Application of the Real-Time Retinex Image Enhancement for Endoscopic Images. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (July 2013), pp. 3407–3410.
- [74] PAN, G., AND WANG, L. Swallowable Wireless Capsule Endoscopy : Progress and Technical Challenges. *Gastroenterol Res Pr.* 2012 (2012).
- [75] QUINLAN, J. R. Induction of decision trees. *Mach Learn* 1, 1 (Mar. 1986), 81–106.
- [76] RAMSOEKH, D., HARINGSMA, J., POLEY, J. W., VAN PUTTEN, P., VAN DEKKEN, H., STEYERBERG, E. W., VAN LEERDAM, M. E., AND KUIPERS, E. J. A Back-to-Back Comparison of White Light Video Endoscopy with Autofluorescence Endoscopy for Adenoma Detection in High-Risk Subjects. *Gut* 59, 6 (June 2010), 785–793.
- [77] REX, D., CUTLER, C., LEMMEL, G., RAHMANI, E., CLARK, D., HELPER, D., LEHMAN, G., AND MARK, D. Colonoscopic Miss Rates of Adenomas Determined by Back-to-Back Colonoscopies. *Gastroenterology* 112, 1 (Jan. 1997), 24–28.
- [78] ROMAIN, O., HISTACE, A., SILVA, J., AYOUB, J., GRANADO, B., PINNA, A., DRAY, X., AND MARTEAU, P. Towards a Multimodal Wireless Video Capsule for Detection of Colonic Polyps as Prevention of Colorectal Cancer. In *13th IEEE International Conference on BioInformatics and BioEngineering* (Nov. 2013), pp. 1–6.
- [79] RONDONOTTI, E., KOULAOUZIDIS, A., KARARGYRIS, A., GIANNAKOU, A., FINI, L., SONCINI, M., PENNAZIO, M., DOUGLAS, S., SHAMS, A., LACHLAN, N., ZAHID, A., MANDELLI, G., AND GIRELLI, C. Utility of 3-Dimensional Image Reconstruction in the Diagnosis of Small-Bowel Masses in Capsule Endoscopy (with Video). *Gastrointest. Endosc.* 80, 4 (Oct. 2014), 642–651.
- [80] RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATY, A., KHOSLA, A., BERNSTEIN, M., BERG, A. C., AND FEI-FEI, L. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis* 115, 3 (Dec. 2015), 211–252.

-
- [81] SAMUEL, A. 1959. some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3, 211–229.
 - [82] SEJNOWSKI, T. J., KIENKER, P. K., AND HINTON, G. E. Learning Symmetry Groups with Hidden Units : Beyond the Perceptron. *Phys. Nonlinear Phenom.* 22, 1 (Oct. 1986), 260–275.
 - [83] SEO, S.-W., AND KIM, M. Efficient architecture for circle detection using Hough transform. In *2015 International Conference on Information and Communication Technology Convergence (ICTC)* (Jeju Island, South Korea, Oct. 2015), IEEE, pp. 570–572.
 - [84] SHANNON, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* 27, 3 (July 1948), 379–423.
 - [85] SHIM, K.-N., JEON, S. R., JANG, H. J., KIM, J., LIM, Y. J., KIM, K. O., SONG, H. J., LEE, H. S., PARK, J. J., KIM, J. H., CHUN, J., PARK, S. J., YANG, D.-H., MIN, Y. W., KEUM, B., AND LEE, B.-I. Quality Indicators for Small Bowel Capsule Endoscopy. *Clin Endosc* 50, 2 (Mar. 2017), 148–160.
 - [86] SPYROU, E., AND IAKOVIDIS, D. K. Video-Based Measurements for Wireless Capsule Endoscope Tracking. *Meas Sci Technol* 25, 1 (Nov. 2013), 015002.
 - [87] SWAIN, P. Wireless Capsule Endoscopy. *Gut* 52, suppl 4 (June 2003), iv48–iv50.
 - [88] SZEGEDY, C., LIU, W., JIA, Y., SERMANET, P., REED, S., ANGUELOV, D., ERHAN, D., VANHOUCKE, V., AND RABINOVICH, A. Going Deeper With Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 1–9.
 - [89] TAJBAKHSI, N., GURUDU, S. R., AND LIANG, J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. *IEEE Trans Med Imaging* 35, 2 (Feb. 2016), 630–644.
 - [90] THAN, T. D., ALICI, G., ZHOU, H., AND LI, W. A Review of Localization Systems for Robotic Endoscopic Capsules. *IEEE Trans Biomed Eng* 59, 9 (Sept. 2012), 2387–2399.
 - [91] VAN GOSSUM, A., MUNOZ-NAVAS, M., FERNANDEZ-URIEN, I., CARRETERO, C., GAY, G., DELVAUX, M., LAPALUS, M. G., PONCHON, T., NEUHAUS, H., PHILIPPER, M., COSTAMAGNA, G., RICCONI, M. E., SPADA, C., PETRUZZIELLO, L., FRASER, C., POSTGATE, A., FITZPATRICK, A., HAGENMULLER, F., KEUCHEL, M., SCHOOF, N., AND DEVIÈRE, J. Capsule Endoscopy versus Colonoscopy for the Detection of Polyps and Cancer. *N Engl J Med* 361, 3 (July 2009), 264–270.
 - [92] WALTHER, J. S. A Unified Algorithm for Elementary Functions. In *Proceedings of the May 18-20, 1971, Spring Joint Computer Conference* (New York, NY, USA, 1971), AFIPS '71 (Spring), ACM, pp. 379–385. event-place : Atlantic City, New Jersey.

- [93] WU, X., AND KUMAR, V. *The Top Ten Algorithms in Data Mining*. CRC Press, Apr. 2009.
- [94] YU, L., CHEN, H., DOU, Q., QIN, J., AND HENG, P. A. Integrating Online and Offline Three-Dimensional Deep Learning for Automated Polyp Detection in Colonoscopy Videos. *IEEE J Biomed Health Inf.* 21, 1 (Jan. 2017), 65–75.
- [95] YUEN, H., PRINCEN, J., ILLINGWORTH, J., AND KITTLER, J. Comparative study of Hough Transform methods for circle finding. *Image and Vision Computing* 8, 1 (Feb. 1990), 71–77.
- [96] ZADEH, L. Probability measures of Fuzzy events. *Journal of Mathematical Analysis and Applications* 23, 2 (Aug. 1968), 421–427.
- [97] ZADEH, L. A., KLIR, G. J., AND YUAN, B. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems : Selected Papers*. World Scientific, 1996. Google-Books-ID : wu0dMiHwJkC.
- [98] ZHANG, R., ZHENG, Y., POON, C. C. Y., SHEN, D., AND LAU, J. Y. W. Polyp Detection during Colonoscopy Using a Regression-Based Convolutional Neural Network with a Tracker. *Pattern Recognit.* 83 (Nov. 2018), 209–219.
- [99] ZHAO, Q., DASSOPOULOS, T., MULLIN, G. E., MENG, M. Q., AND KUMAR, R. A Decision Fusion Strategy for Polyp Detection in Capsule Endoscopy. *Stud Health Technol Inf.* 173 (2012), 559–565.
- [100] ZHOU, X., ITO, Y., AND NAKANO, K. 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* (2014), IEEE, pp. 447–452.