

La détection des polypes dans les images issues de coloscopies

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

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

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

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

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

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

2.1.1 Constat sur la forme des polypes

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



FIGURE 2.1 – Illustration concernant la forme des polypes

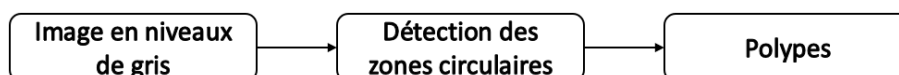


FIGURE 2.2 – Chaîne de traitement pour la détection des polypes basée sur la forme

2.1.2 La détection des formes circulaires

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

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

La transformée de Hough

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

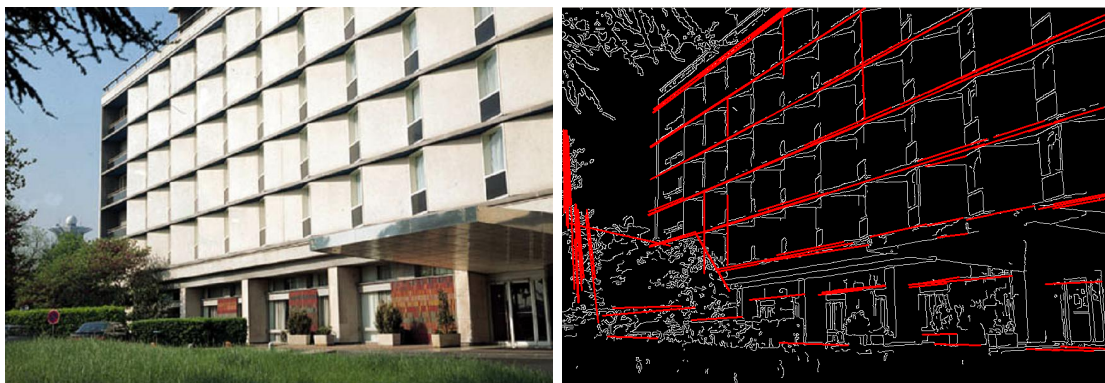


FIGURE 2.3 – Illustration de la transformée de Hough (Source : OpenCV Docs)

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

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

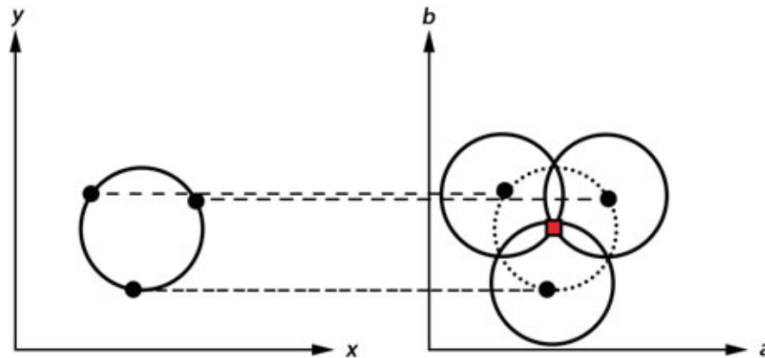


FIGURE 2.4 – Illustration de la transformée de Hough circulaire. Sur la gauche, trois motifs circulaires dans l'espace de l'image, et sur la droite, trois cercles correspondants au point rouge dans l'espace paramétrique (de coordonnées (a, b) pour un rayon $\mathcal{R}$ fixé)

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

La méthode des filtres Log-Gabor

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

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

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

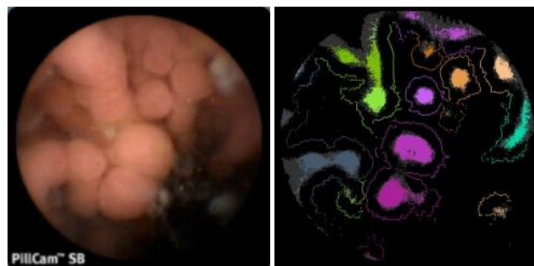


FIGURE 2.5 – Résultat de la détection basée sur les filtres Log-Gabor [KB09]

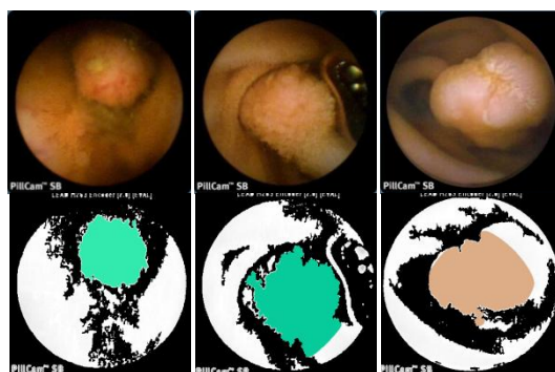


FIGURE 2.6 – Résultat de la détection basée sur les filtres Log-Gabor [KB09]

Conclusion intermédiaire sur la méthode à adopter

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

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

- Les faux négatifs (False Negative, FN). Une détection "fausse négative" est le résultat lorsqu'un polype est signalé positif alors qu'il n'y a rien.
- Le **rappel** ($= \frac{TP}{TP+FN}$). Le rappel permet de définir le résultat de bonnes détections par rapport au nombre total de lésions à détecter.
- La **précision** ($= \frac{TP}{TP+FP}$). La précision permet de définir le résultats de bonnes détections par rapport au nombre total de lésions détectées.
- Le **F₁ Score** ($= \frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$). Le F₁ Score, qui combine Précision et Rappel, est utilisé pour faire la balance entre le nombre de fausses alarmes et le nombre de lésions ratées.

Méthode de détection	Rappel	Précision	F ₁ Score
Transformée de Hough	94%	14%	24%
Log Gabor	42%	77%	54%

TABLE 2.1 – Performances de la détection des polypes

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

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

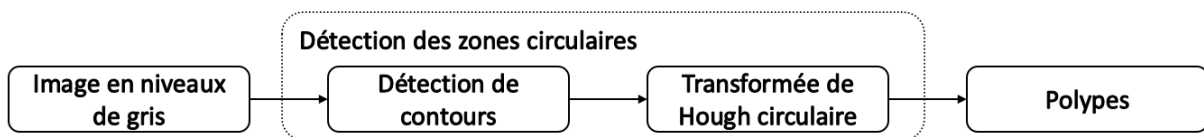


FIGURE 2.7 – Chaîne de traitement pour la détection des polypes basée sur la transformée de Hough

Limitation de la transformée de Hough pour la détection des polypes

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

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

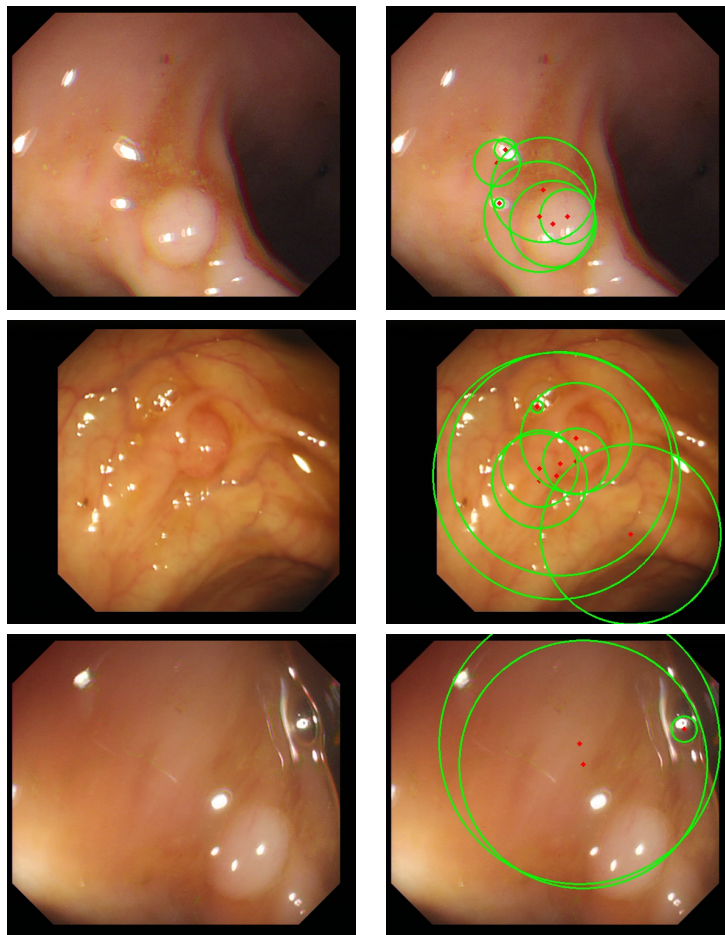


FIGURE 2.8 – Exemples de détections avec la transformée de Hough circulaire. Les images de droite montrent le résultat de la transformée de Hough sur les images originales correspondantes à gauche.

2.1.3 Apprentissage machine en complément

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

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

Une chaîne de traitement en complément de la transformée de Hough

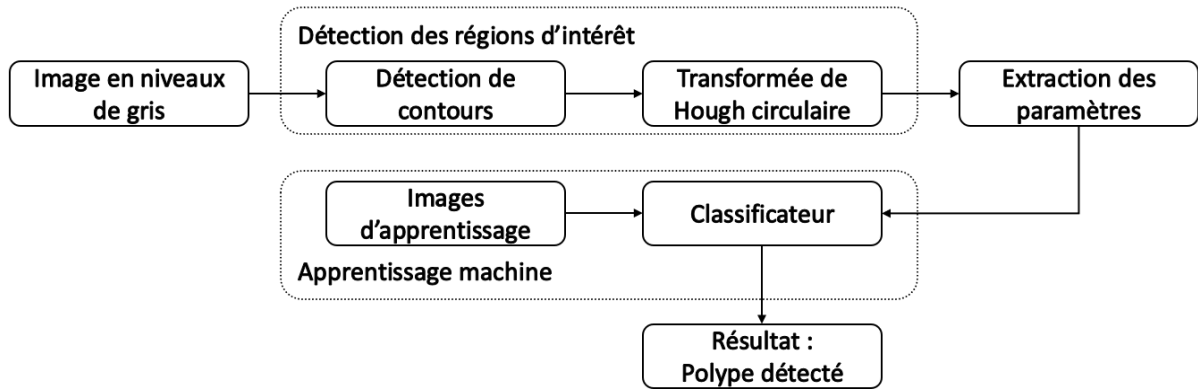


FIGURE 2.9 – Schéma de détection des polypes avec de l'apprentissage machine

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

Matrice de cooccurrence et paramètres d'Haralick

Matrice de cooccurrence

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

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

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

FIGURE 2.10 – Illustration du calcul de la matrice de cooccurrence

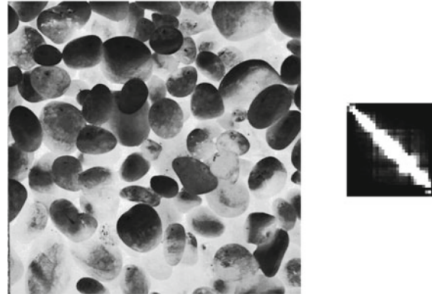


FIGURE 2.11 – Exemple d’une matrice de cooccurrence associée à son image

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

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

Paramètres d’Haralick

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

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

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

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

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

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

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

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

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

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

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

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

FIGURE 2.12 – Exemples de paramètres d'Haralick

Performances de l'apprentissage machine

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

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

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

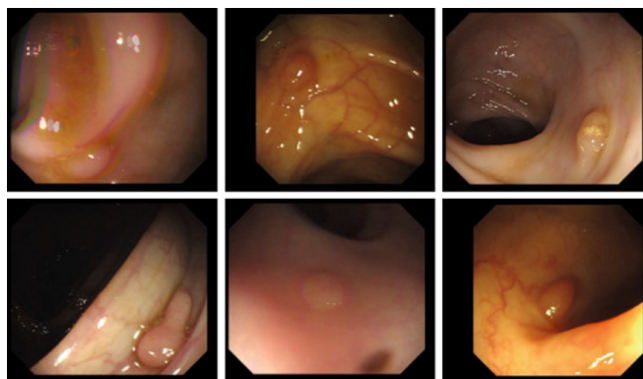


FIGURE 2.13 – Exemples d’images disponibles dans la base CVC-Colon

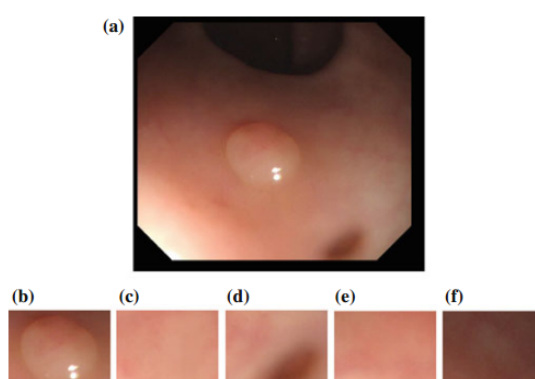


FIGURE 2.14 – Exemple de création des patches : (a) image originale, (b) exemple positif, (c-f) exemples négatifs

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

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

- Le rappel
- La précision
- Le F_1 Score

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

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

Méthode d'apprentissage	Rappel	Précision	F ₁ Score
LVQ	92%	62%	74%
Adaboost	77%	72%	74.4%
Cascade Adaboost	91%	83%	86.1%

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

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

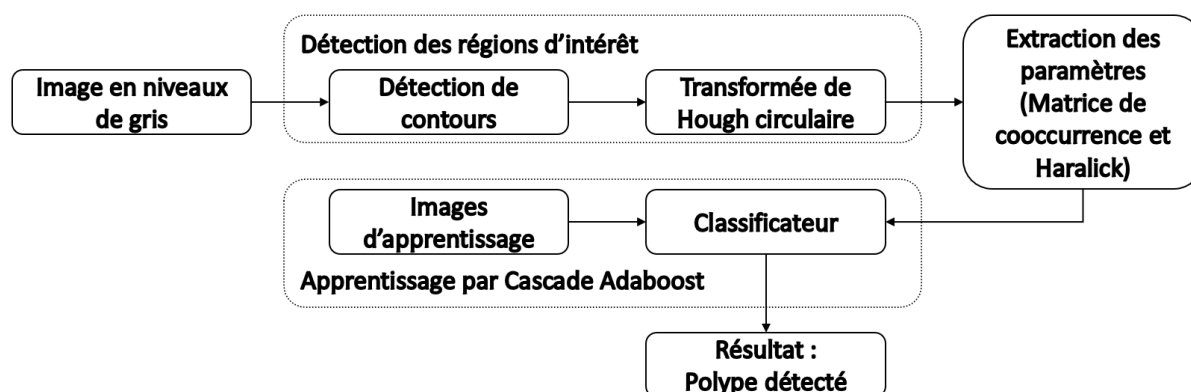


FIGURE 2.15 – Schéma de détection des polypes avec de l'apprentissage Cascade Adaboost

2.1.4 Conclusion et limitations de cette méthode

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

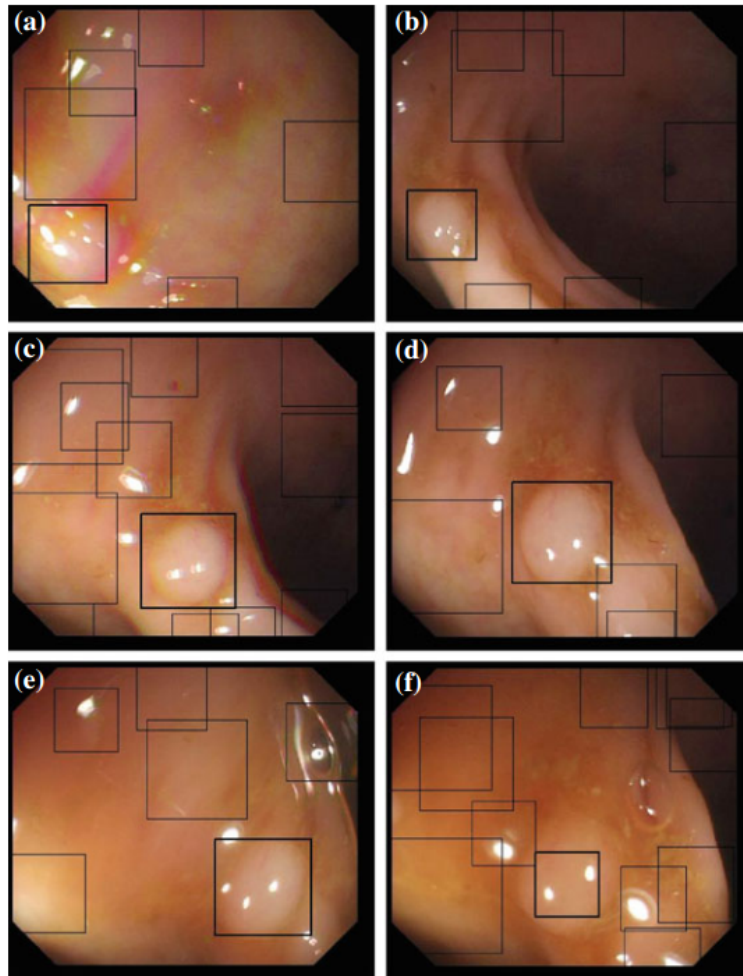


FIGURE 2.16 – Exemples de détections de polypes dans des images, les rectangles noirs sont les zones détectées par la transformée de Hough, les rectangles noirs plus épais sont les zones classées comme polype après l'étape de classification

machine de travail qui possède une configuration très performante.

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

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

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

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

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

2.2 L'apprentissage actif pour la détection des polypes

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

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

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

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

2.2.1 Implémentation matérielle

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

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

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

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

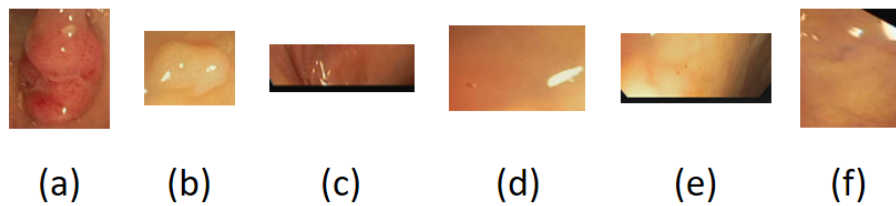


FIGURE 2.17 – Imageriettes pour la préparation de l'apprentissage, (a) (b) représentent des imageriettes avec polypes, (c) (d) (e) (f) représentent des imageriettes sans polypes

2.2.2 Paramètres d'apprentissage

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

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

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

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

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

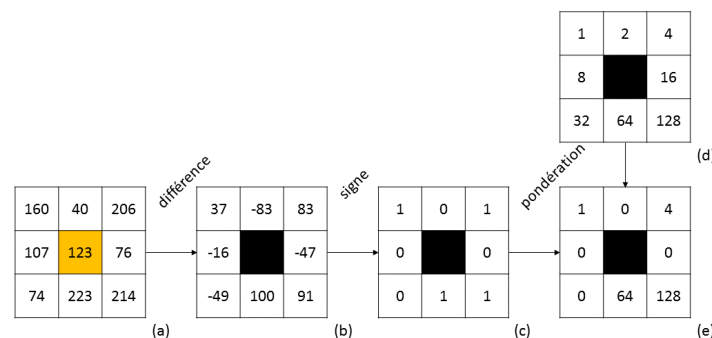


FIGURE 2.18 – Exemple d'utilisation des LBP

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

Paramètres pseudo-Haar

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



FIGURE 2.19 – Exemples de masques proposés par Viola et Jones

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

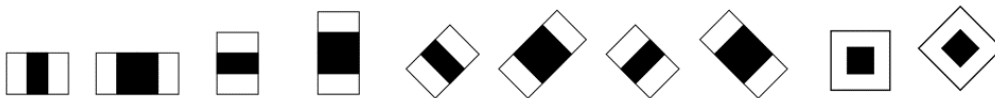


FIGURE 2.20 – Exemples de masques proposés par Lienhart et Maydt

2.2.3 L'algorithme de base

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

2.2.4 Résultats préliminaires sur la méthode de détection

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

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

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

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

- Les vrais positifs (True Positive, TP). Une détection "vraie positive" est le résultat lorsqu'un polype est détecté correctement.
- Les faux positifs (False Positive, FP). Une détection "fausse positive" est le résultat lorsqu'un polype est signalé négatif alors qu'il devrait être positif.
- Les faux négatifs (False Negative, FN). Une détection "fausse négative" est le résultat lorsqu'un polype est signalé positif alors qu'il n'y a rien.
- Le rappel ($= \frac{TP}{TP+FN}$). Le rappel permet de définir le résultat de bonnes détections par rapport au nombre total de lésions à détecter.
- La précision ($= \frac{TP}{TP+FP}$). La précision permet de définir le résultats de bonnes détections par rapport au nombre total de lésions détectées.
- Le F₁ Score ($= \frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$). Le F₁ Score, qui combine Précision et Rappel, est utilisé pour faire la balance entre le nombre de fausses alarmes et le nombre de lésions ratées.
- Le **temps de détection moyen pour une image**. Il s'agit du temps moyen (moyenne prise sur l'ensemble des 273 images) mis par la méthode pour effectuer la détection des polypes dans une image.

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

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

TABLE 2.4 – Résultats de la détection sur images fixes

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

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

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

Principe de l'apprentissage actif

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

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

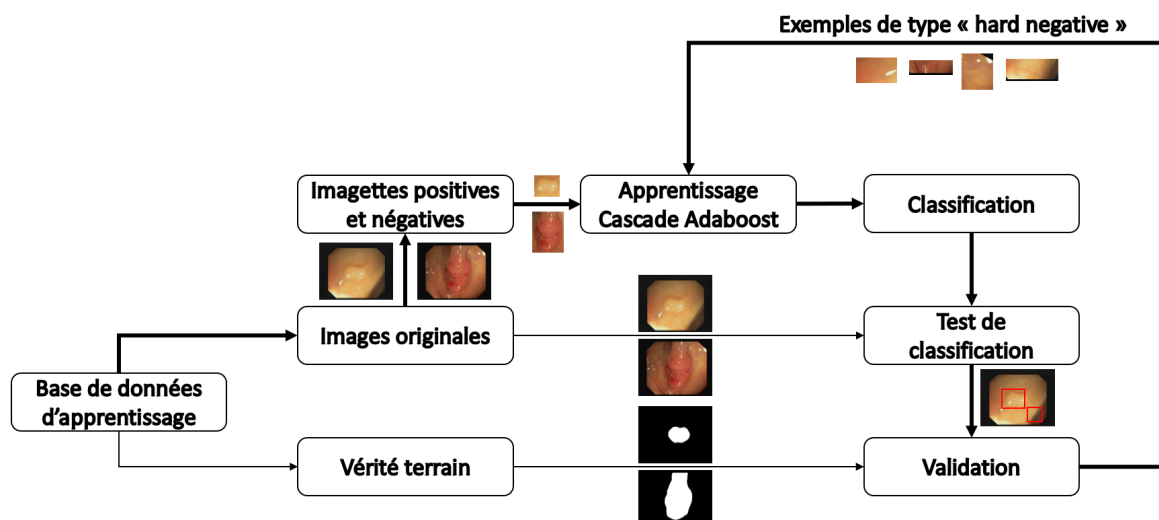


FIGURE 2.21 – Schéma de principe de l'apprentissage actif

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

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

TABLE 2.5 – Principe de l'apprentissage actif

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

Effets de l'apprentissage actif sur les performances

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

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

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

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

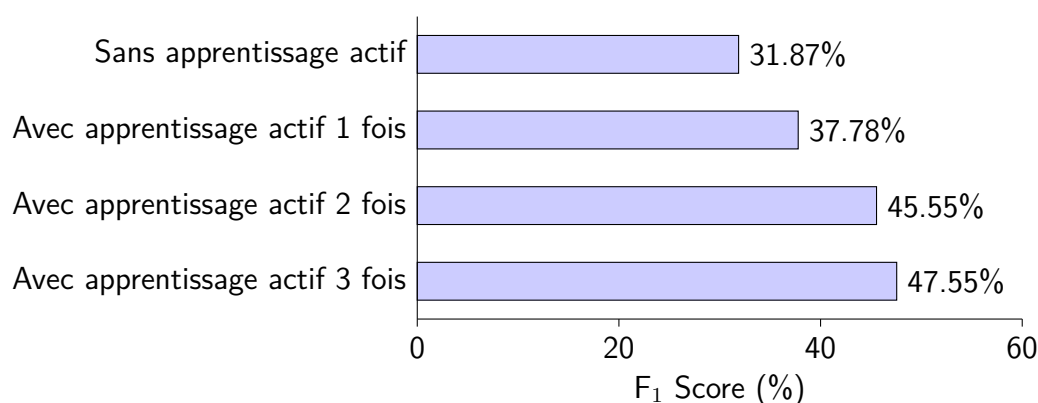


FIGURE 2.22 – Évolution du F_1 Score avec les répétitions d'apprentissage actif

Conclusion sur l'apprentissage actif

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

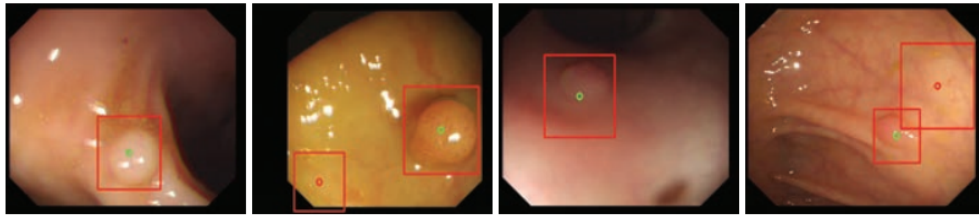


FIGURE 2.23 – Résultats de la méthode avec apprentissage actif

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

2.3 Conclusion sur la détection dans les images

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

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

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

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

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

