

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université 20 Août 1955- Skikda
Faculté des sciences - Département d'informatique



Mémoire en vue de l'obtention du diplôme de master académique
'Réseaux et Systèmes Distribués (RSD)'

*Apprentissage profond pour la détection et la localisation
automatisées des polypes en coloscopie par l'utilisation de la
méthode YOLOv8*

Réalisé par :

TALAA Asma

Dirigé par :

Pr. REDJIMI Mohammed

Année Universitaire 2022-2023

Remerciement

En premier lieu, j'exprime mes remerciements à ALLAH le tout-puissant qui m'a donné le courage et la volonté pour réaliser ce modeste travail.

Je remercie mon enseignant et encadreur de mémoire

***Pr. REDJIMI Mohammed** qui a accepté la direction de ce travail ainsi que pour ses judicieux et précieux conseils, ses compétences scientifiques, sa confiance qu'il m'accordée surtout pour ses très grandes qualités humaines et sa gentillesse.*

Mes remerciements vont aux membres du jury qui ont accepté de juger ce travail.

Enfin, je tiens à exprimer ma profonde reconnaissance à tous les responsables et les enseignants du département d'informatique de

L'université de Skikda

Qui ont contribué ma formation.



ASMA TALAA

Dédicaces

*Je dédie ce travail, fruit de nombreuses années d'études à :
mes chers parents pour leur patience,*

*A la femme qui a souffert sans me laisser souffrir, qui n'a jamais dit non à mes exigences et qui n'a épargné aucun effort pour me rendre heureuse. Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes côtés a toujours été ma source de force pour affronter les différents obstacles :
mon adorable mère.*

*A l'homme, ma précieuse offre de dieu, vous avez toujours été à mes côtés pour me soutenir et m'encourager. Que ce travail reflète ma gratitude et mon amour pour toi :
mon cher père.*

À tous mes frères et mes sœurs pour leur aide, présence et réconfort.

À la joie de ma vie Malek, Maria, Iyed.

À Tous mes enseignants.

À Tous mes amis

À toutes les personnes que j'aime et qui m'aiment

ملخص

يمكن أن تؤثر العديد من الأمراض على الجهاز الهضمي للإنسان. من بينها سرطان القولون والمستقيم، وهو أحد أكبر أسباب الوفيات المرتبطة بالسرطان في العالم. يحدث CRC بشكل رئيسي بسبب الاورام الحميدة التي يمكن أن تنمو على الجدران الداخلية للأمعاء أو المستقيم. لتحسين فرص البقاء على قيد الحياة، يعتبر الاكتشاف المبكر والتشخيص والعلاج في الوقت المناسب من العوامل الحاسمة (يتعلق الأمر بالإزالة الجراحية للزوائد اللحمية). يساهم استخدام أنظمة التشخيص بمساعدة الكمبيوتر (CADx) المتطورة والمجهزة بخوارزميات التعلم الآلي المناسبة، وأساليب التعلم العميق على وجه التحديد في مساعدة الاختصاصيين في الحصول على الكشف عن الأشياء ذات الصلة جدًا بالنتائج في الطب. في هذا الصدد، تعرض هذه المقالة نموذجًا تعليميًا عميقًا للكشف الآلي عن الزوائد اللحمية باستخدام *You Only Look One* (YOLO). تم تجربة الإصدارات السابقة من عائلة YOLO مع تحسين الأداء؛ تتعلق الدراسة الحالية بأحدث إصدار من YOLOv8، والذي لم يتم تضمينه بعد في اكتشاف CRC. يتم استخدام نموذج الكشف في الوقت الفعلي مع مجموعات البيانات Kvasir-SEG و BKAI-IGH NeoPolyp-Small. تتم مقارنة النتائج التي تم الحصول عليها بالنتائج التي تم الحصول عليها باستخدام YOLOv7 وتظهر أداءً حسنًا.

الكلمات المفتاحية: الاورام الحميدة، سرطان القولون والمستقيم (CRC)، التعلم العميق، أنظمة التشخيص بمساعدة الكمبيوتر (CAD)، تنظير القولون، YOLO

Résumé

Plusieurs maladies peuvent affecter le système digestif humain. Parmi elle, le cancer colorectal (CRC) qui présente l'une des plus grandes causes de décès liés au cancer dans le monde. Le CRC est principalement causé par des polypes qui peuvent se développer sur les parois internes des intestins ou du rectum. Pour améliorer les chances de survie, la détection préalable, le pronostic et le traitement rapide sont des facteurs cruciaux. Détecter de façon précoce, l'ablation chirurgicale des polypes peut aider à les éradiquer. L'utilisation de systèmes développés d'outils d'aide au diagnostic assisté par ordinateur (CADx) équipés d'algorithmes d'apprentissage automatique appropriés, et plus particulièrement de méthodes d'apprentissage en profondeur, contribue à aider les praticiens à obtenir une détection d'objets très pertinente d'anomalies en médecine. À cet égard, ce travail expose un modèle d'apprentissage en profondeur pour la détection automatisée des polypes en utilisant You Only Look Once (YOLO) version 8. Les versions antérieures de la famille YOLO ont été expérimentées et ont montré de bonnes performances. Dans ce travail, nous avons utilisé la YOLOv8 qui n'a pas encore été impliquée dans la détection des polypes, à notre connaissance. Le modèle de détection en temps réel est utilisé avec les deux bases de données Kvasir-SEG et BKAI-IGH. Les résultats sont comparés à ceux obtenus par l'utilisation de YOLOv7 et montrent une nette amélioration des performances de YOLOv8 par rapport à YOLOv7.

Mots-clés : Polypes, You Only Look Once (YOLO), Cancer colorectal (CRC), Deep-learning, Systèmes de diagnostic assistés par ordinateur (CADx), coloscopie.

Abstract

Several diseases can affect the digestive human system. Among them, colorectal cancer (CRC) is the great common cause of cancer-related death in the world. Polyps that can grow on the inner walls of the intestines or the rectum mainly cause the CRC. To improve chances of survival, prior detection, prognosis, and timely treatment are crucial factors (it is about the surgical removal of polyps). Using developed Computer-aided diagnosis systems (CADx) systems equipped with appropriate machine learning algorithms, specifically deep learning methods contribute to assisting patricians to obtain very pertinent object detection of abnormalities in medicine. In this respect, this study exposes a deep learning model for automated polyp detection by using the You Only Look Once (YOLO). Earlier versions of the YOLO family have been experimented with and given improved performances; the present paper concerns the latest YOLOv8 version, which has not yet been implicated for CRC detection. The real-time detection model is used with the Kvasir-SEG dataset. The obtained results are compared to those obtained by the use of the YOLOv7 and show improved performances.

Keywords: *Polyps, You Only Look Once (YOLO), Colorectal Cancer (CRC), Deep-learning, Computer-aided diagnosis systems (CADx), colonoscopy*

Table des matières

Remerciement.....	2
Dédicaces	3
Résumé.....	5
Table des matières.....	7
Liste des abréviations.....	9
Liste des tableaux.....	11
Table des figures.....	12
Introduction générale.....	12
Chapitre1 : Généralités et état de l’art.....	17
1. Introduction.....	17
2. Généralités sur le tube digestif humain.....	17
2.1. Les méthodes exploratrices du tube digestif humain.....	18
3. Le cancer colorectal.....	20
4. Détection et localisation des polypes.....	23
5. Apports de l’intelligence artificielle.....	23
6. Notion de traitement des images côloscopies.....	24
1. Système d’IA.....	24
7. Systèmes de coloscopie assistés par l’IA.....	25
7.1. Détection assistée par ordinateur (CAD) des polypes.....	25
7.2. Classification des polypes.....	25
7.3. Aide à la navigation.....	26
8. Recherche et études cliniques.....	26
9. Les défis et les limites associés à l’intégration des technologies de l’IA dans la coloscopie.....	27
10. Impacts potentiels de l’IA en coloscopie.....	28
11. Conclusion.....	30
Chapitre 02: L’apprentissage profond (Deep-Learning: DL).....	31
1. Introduction.....	31
2. L’intelligence artificielle.....	31

3. L'Apprentissage automatique	32
3.1. L'apprentissage supervisé	32
3.2. L'apprentissage non supervisé	33
3.3. L'apprentissage par renforcement	33
4. L'apprentissage profond.....	33
4.1. Les réseaux de neurones profonds	34
4.2. Les réseaux de neurones convolutifs.....	34
5. Vision par ordinateur.....	39
5.1. La détection en deux étapes	40
5.2. La détection en une étape	43
6. Conclusion.....	49
Chapitre 03 : L'approche proposée	50
1. Introduction	50
2. Le modèle YOLOv8.....	50
2.1. Architecture de YOLOv8	51
3. Les Datasets.....	52
4. Approche proposée.....	55
4.1. Prétraitement des données	55
4.2. Apprentissage de modèle YOLOv8	59
4.3. Métriques d'évaluation des performances	61
5. Conclusion.....	63
Chapitre 04 : L'implémentation et les résultats obtenus	64
1. Introduction	64
2. Le matériel utilisé.....	64
3. Environnements et outils de développement utilisés	64
4. L'implémentation	68
4.1. Les interfaces graphiques de l'application	68
5. Les résultats obtenus	71
5.1. Apprentissage	71
5.2. Evaluation du modèle.....	74
5.3. Test et résultats.....	76
6. Etudes comparatives.....	78
7. Conclusion.....	79
Références	81

Liste des abréviations

CADx : Computer-aided diagnosis systems

CNN: Convolutional neural network

CRC: Colorectal Cancer

CSP: Cross Stage Partial

CV: Computer Vision

DL: Deep Learning

DNN: Deep Neural Networks

FC: Fully Connected

FCN: Fully Convolutional Network

Fps: Frames Per Seconde

GPU: Graphical Processing Unit

IA: Intelligence Artificiel

IOU: Intersection Over Union

IRM: Imagerie Par Resonance Magnétique

ML: Machine Learning

MLP: Multi Layers Perceptron

NMS: Non-Maximum Suppression

R-CNN: Region Proposal with Convolutional Neural Network

ReLU: Rectified Linear Unit

ROI: Region of Interest

RPN: Region Proposal Network

SiLU: Sigmoid-weighted Linear Unit

SPP-Net: Spatial Pyramid Pooling Network

SSD: Single Shot Detector

SVM: Support Vector Machine

YOLO: You Only look Once

Liste des tableaux

Tableau 1 - Datasets endoscopique disponible.	53
Tableau 2 - exemple d'augmentation data.	58
Tableau 3 - comparaison entre les vrai et les faux positive et négative.....	61
Tableau 4- Résultats des performances du modèle YOLOv8.....	74
Tableau 5- Résultats de comparaison entre YOLOv8 et YOLOv7.	79

Table des figures

Figure 1 - Le tube digestif humain [2]	18
Figure 2 - Capsule vidéo-endoscopique [7]	19
Figure 3 - Illustration du côlon humain [9]	20
Figure 4 - Illustration concernant différentes formes de polypes [9]	21
Figure 5 - Classification morphologique des polypes [9]	21
Figure 6 - illustration de l'évolution d'un polype vers une tumeur [10]	22
Figure 7 - illustration quelques exemples de détection des polypes. [11]	23
Figure 8 - comparaison entre les neurones biologiques et le neurone artificiel [12]	24
Figure 9 - Système d'IA pour la détection et la localisation de polypes	24
Figure 10 - Hiérarchie des domaines de l'intelligence artificielle [27]	31
Figure 11- Architecture d'apprentissage automatique [29]	32
Figure 12 - Architecture d'apprentissage profond [34]	34
Figure 13 - Architecture d'un réseau de neurones convolutive [38]	35
Figure 14 - La couche de convolution [25]	36
Figure 15- le Fonctionnement de la couche Pooling [27]	37
Figure 16 - Fonction d'activation ReLU [29]	38
Figure 17- Couche entièrement connectée [27]	38
Figure 18 - Tâches de vision par ordinateur [32]	39
Figure 19 - Structure typique du détecteur d'objets [33]	40
Figure 20 - Approche en deux étapes	40
Figure 21- Architecture de modèle R-CNN [24]	41
Figure 22- architecture de model Fast R-CNN [25]	42
Figure 23 - Architecture de modèle Faster R-CNN [38]	43
Figure 24 - Architecture de base du modèle YOLO [50]	44
Figure 25 - Schéma du fonctionnement de YOLO [51]	44
Figure 26 - boîte englobant [51]	45
Figure 27- Intersection de la boîte englobante [54]	46
Figure 28 - Le résultat de Suppression non maximale [52]	47
Figure 29 - Une chronologie des versions de YOLO [52]	47

Figure 30 - YOLOv8 par rapport aux autres modèles YOLO [58]	51
Figure 31- Architecture YOLOv8 [51]	51
Figure 32 - Nouveau module YOLOv8 C2f.....	52
Figure 34 - Polype coloscopique, boîtes englobantes correspondantes et masques de segmentation issus de Kvasir-SEG.[52].....	54
Figure 35 - Polype coloscopique, boîtes englobantes correspondantes et masques de segmentation provenant de BKAI-IGH NeoPolyp-Small [53]	54
Figure 36 - Présentation de la méthodologie de la tâche de détection et de localisation des polypes	55
Figure 37 - données combinées pour entraîner le système de détection de polypes.....	55
Figure 38 - exemple annotation utilisant roboflow.	56
Figure 39 - exemple fichier d'annotation de la base de données.	56
Figure 40 - Contenu du fichier polyp.yaml pour YOLOv8.....	59
Figure 41- logo de Google colab	64
Figure 42 - L'environnement Google Colaboratory	65
Figure 43 - Logo de python.....	66
Figure 44 - Logo de roboflow.	66
Figure 45 - Logo OpenCV	67
Figure 46 - Logo tkinter.	67
Figure 47 - Logo PyCharm.....	68
Figure 48 - L'interface principale de l'application.	69
Figure 49 - Les tâches de l'application.	69
Figure 50 - Le choix d'une image.	70
Figure 51 - Le choix d'un enregistrement vidéo.	71
Figure 52 - Résumé du modèle YOLOv8 lors de l'entraînement du modèle.	71
Figure 53 - Graphique des performances de l'entraînement du modèle YOLOv8.....	72
Figure 54 - la courbe de précision. Figure 55 - La courbe du rappel.	73
Figure 56 - La courbe PR (Precision-Recall) pour l'entraînement du modèle.	74
Figure 57- la matrice de confusion de la classe polype.....	75
Figure 58 - Étiquettes et prédictions correspondantes pour la validation par lots (batch).	76
Figure 59 - Quelques résultats de tests de détection obtenus par YOLOv8.....	77
Figure 60 - Détection des polypes dans une vidéo par YOLOv8.....	78

Introduction générale

Contexte

La médecine a depuis longtemps profité des avancées technologiques pour l'aide à la détection, à l'analyse et au diagnostic des maladies. Depuis le simple stéthoscope qui est devenu l'emblème des médecins et qui représente un peu une image d'Épinal de cette noble profession aux puissants algorithmes d'intelligence artificielle couplés à des procédés sophistiqués d'IRM (Imagerie par Résonance Magnétique), la technologie est utilisée à bon escient et accompagne les actes, raisonnements et décision des médecins dans les traitements quotidiens des patients. Certaines maladies malignes et dévastatrices dont les effets peuvent être irréversibles se développent dans un silence absolu en ne causant aucune douleur ni signe avant-coureur et peuvent aboutir à des cancers. Si les causes sont détectées de façon précoce, de simples actes médicaux peuvent pallier le développement de ces maladies. Le problème est encore plus inquiétant lorsqu'il s'agit de maladies internes qui nécessitent des examens et des outillage particuliers, très souvent très onéreux et sophistiqués, et dont les investigations sont parfois invasives. Ce travail se place dans ce contexte et concerne l'examen du colon en vue de la détection et l'élimination précoces des polypes pour prévenir le cancer colorectal (CRC) qui représente l'une des causes principales de nombreux décès dans le monde [1]. Les polypes qui sont la croissance anormale de tissus à l'intérieur du côlon peuvent entraîner un cancer colorectal [2] s'ils ne sont pas détectés et soignés à temps. Dans ce contexte, la coloscopie [3] est la méthode la plus utilisée pour examiner le côlon et détecter l'existence ou l'absence de polypes. C'est une méthode souvent très invasive et pratiquée, selon l'avis du patient, avec ou sans anesthésie. La coloscopie nécessite une préparation intestinale de qualité et une insufflation d'air lors de l'examen intestinal [4]. C'est une technique onéreuse en temps et dont les coûts économiques sont assez élevés. De plus, cette technique nécessite une attention particulière de la part du gastro-entérologue et l'expérience montre que les polypes (surtout ceux de petites tailles) ne sont pas très souvent détectés lors de l'examen et peuvent causer des dommages irréversibles dans le futur. Dans la méthode classique, appelée endoscopie, le gastro-entérologue insère un long tube flexible dans le rectum du patient pour observer l'intérieur du côlon et détecter d'éventuelles anomalies.

Objectifs

Ce travail vise à proposer une approche utilisant l'intelligence artificielle et plus particulièrement l'apprentissage profond (deep-learning) pour la détection des polypes dans des séquences d'images obtenues par examen l'examen des parois internes du colon. Les algorithmes mis en œuvre doivent concilier l'optimisation de la rapidité de la détection, les moyens de traitement informatique et les procédés de stockage et de consultation des bases de données dédiées (data sets). Il est à noter que le traitement d'images numériques se heurte au poids des images qui sont composées de pixels et dont la taille devient vite exorbitante ce qui induit une complexité exponentielle en temps de traitement et en capacité de stockage.

Méthodologie

L'approche proposée s'inspire et repose sur l'utilisation des réseaux de neurones convolutifs (CNN : Convolutional Neural Networks). Une application pratique a été développée par l'utilisation de YOLOv8 (You Only Look Once) qui est, à l'heure actuelle la dernière version de YOLO. La mise en œuvre de cette application est effectuée en Python et l'utilisation de la combinaison des deux bases de données KVASIR-SEG et BKAI-IGH NeoPolyp-Small qui sont gratuitement accessibles et téléchargeables via le net.

Résultats

Les résultats obtenus sont encourageants et montrent une précision qui se situe aux environs de 96% lorsque l'on utilise le modèle YOLOv8 alors qu'avec YOLOv7, les chiffres plafonnent à 86,54%. Ceci étant dû, d'une part à ce que YOLOv8 intègre des fonctionnalités beaucoup plus étendues que YOLOv7 et aussi au fait de la combinaison des deux bases de données KVASIR-SEG et BKAI-IGH NeoPolyp-Small ce qui a pour but d'augmenter la taille de la base de données (en la multipliant par deux pratiquement). Enfin la troisième amélioration réside dans l'augmentation des données en intégrant de nouvelles images à partir des images contenues dans les datasets (Data Augmentation). Les résultats de ce travail ont fait l'objet d'une communication nationale présentée à la première conférence nationale informatique et application (RIA'23) tenue à l'université de Relizane le 10 Mai 2023.

Structure du mémoire

Le mémoire est organisé comme suit :

- Le premier chapitre introduit les notions de base concernant la coloscopie et les matériels et appareils utilisés ainsi que l'apport de l'intelligence artificiel
- Le second chapitre est consacré aux approches de l'intelligence artificielle et particulièrement du deep-learning ainsi que l'utilisation du modèle YOLO dans la détection et la localisation d'objets.
- Le troisième chapitre présente le modèle YOLOv8 et son application dans la détection des polypes. La version 8 est apparue en janvier 2023 et c'est cette version qui est celle utilisée dans ce travail.
- Le quatrième et dernier chapitre décrit l'implémentation du système, l'application obtenue et quelques interfaces ainsi que les résultats et leur discussion.
- Une conclusion générale et quelques perspectives clôturent ce travail.

Chapitre 1 : état de l'art

1. Introduction

Le cancer colorectal est l'un des cancers les plus répandus dans le monde, avec une incidence en constante augmentation. Malheureusement, la détection précoce de cette maladie reste un défi important, en grande partie en raison de la difficulté à visualiser et à caractériser les lésions précancéreuses lors des examens de coloscopie. C'est là que l'intelligence artificielle intervient, offrant des solutions innovantes pour améliorer la détection et la classification des lésions lors des coloscopies.

Dans ce premier chapitre, nous allons explorer différentes méthodes exploratrices du tube digestif humain, en mettant l'accent sur les techniques de coloscopie assistées par l'IA. Nous allons également examiner les apports de l'intelligence artificielle dans le domaine de la coloscopie, ainsi que les défis et les limites associés à l'intégration de ces technologies dans la pratique clinique.

2. Généralités sur le tube digestif humain

Le tube digestif humain est un appareil très complexe qui nous permet de nous alimenter et dont la structure est présentée figure 1. Plusieurs maladies peuvent attaquer ce tube digestif dont plusieurs ne donnent aucun signe permettant de les déceler surtout que la quasi-totalité de ce système est enfouie dans le corps humain et qu'il s'étend sur neuf mètres environ dont 7 à 8 m pour l'intestin grêle. Afin de déceler les éventuelles maladies, le gastro-entérologue utilise des outils et matériels souvent très sophistiqués pour explorer l'intérieur du tube digestif. [1]

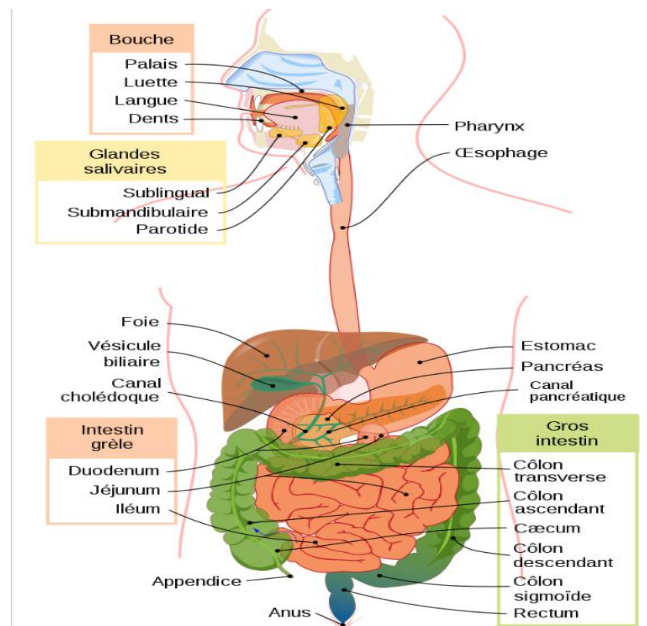


Figure 1 - Le tube digestif humain [2]

2.1. Les méthodes exploratrices du tube digestif humain

Afin d'explorer l'intérieur du tube digestif, le praticien a recours à l'endoscopie ou fibroscope qui est une méthode d'observation visuelle de l'intérieur cavité du tube inaccessible à l'œil. Le premier instrument utilisé, et qui est toujours d'usage est appelé endoscope [3] ou fibroscope [4]. Cet instrument est composé d'un long tube optique muni d'un système d'éclairage et, éventuellement, d'une caméra vidéo qui permet de retransmettre les images sur un écran.

On distingue deux types de fibroscopie : la fibroscopie haute qui concerne la partie allant de la bouche à l'estomac et la fibroscopie basse qui concerne le reste du tube digestif. La fibroscopie est une méthode invasive et elle est pratiquée sous ou sans anesthésie locale ou générale selon le vœu du patient.

- **La coloscopie** : la coloscopie permet d'explorer le côlon jusqu'à la jonction avec l'intestin grêle, ce qui fait environ une longueur d'un mètre et demi. Elle est réalisée après régime sans résidu et après lavement colique total. Elle est généralement réalisée sous anesthésie générale. La coloscopie est indiquée en complément de l'examen clinique pour le dépistage sur facteurs de risque des tumeurs coliques bénignes (polypes, adénomes) et malignes (cancer du côlon) ou encore d'affections inflammatoires de l'intestin (maladie de Crohn, rectocolite hémorragique). Elle permet enfin la réalisation de gestes à visée diagnostique (biopsies) ou thérapeutique (ablation

de polypes...). La coloscopie comporte des risques (lésion de la paroi colique, mauvaise tolérance de l'anesthésie générale, etc.). [5]

Aujourd'hui avec les progrès et avancées des nouvelles technologies, plusieurs techniques exploratrices récentes sont utilisées :

- **Les capsules vidéo-endoscopiques** : Une capsule vidéo-endoscopique (endoscopie avec caméra sans fil) est une petite capsule miniature munie d'une unité d'alimentation en énergie électrique (pile), de petites caméras, d'une source lumineuse et d'une antenne pour la communication d'informations. Le principe consiste à faire avaler au patient une capsule qui va suivre tout le tube digestif et envoyer des milliers d'images de ses parois intérieures. Ces images de la muqueuse intestinale sont transmises à des capteurs portés à proximité par le patient (au niveau de la ceinture ou d'une poche). Cette technique est moyennement invasive.

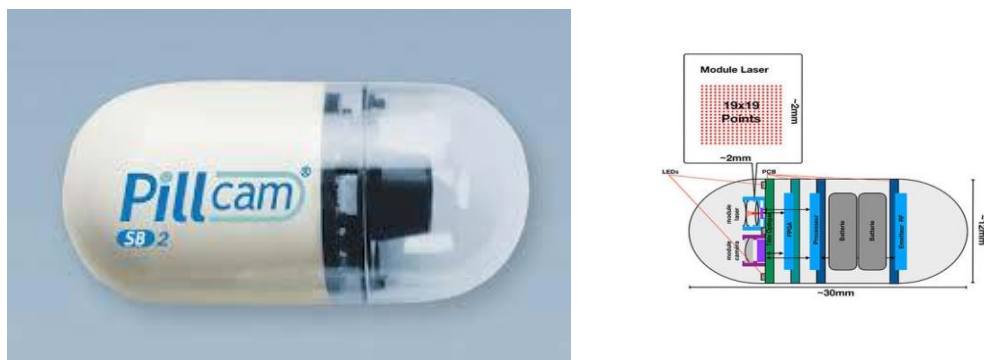


Figure 2 - Capsule vidéo-endoscopique [7]

La capsule vidéo-endoscopique est particulièrement indiquée pour découvrir un saignement digestif et des problèmes au niveau de la muqueuse de l'intestin grêle, qui est une zone difficile d'accès avec un endoscope. La capsule est en général évacuée dans les selles dans un délai de 12 heures environ, parfois sans que cela soit remarqué. Il n'est pas nécessaire de retirer la capsule et elle peut être éliminée dans les toilettes. Si la personne ne remarque pas l'évacuation de la capsule, les médecins peuvent réaliser des radiographies ou une tomodensitométrie pour déterminer si elle est toujours présente dans le tube digestif. Dans de rares cas, la capsule peut être coincée dans le tube digestif et les médecins peuvent avoir besoin de réaliser une endoscopie ou une intervention chirurgicale pour la retirer.[6]

- **L'échographie** : la tomodensitométrie (TDM) et l'imagerie par résonance magnétique (IRM) du tube digestif sont des techniques non-invasives d'exploration du tube digestif. Leur principe réside dans l'obtention d'images radiographiques de l'intérieur du tube digestif. Cet examen génère des images bi- et tridimensionnelles du côlon qui peuvent mettre en évidence des tumeurs ou d'autres problèmes. [8]

3. Le cancer colorectal

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

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

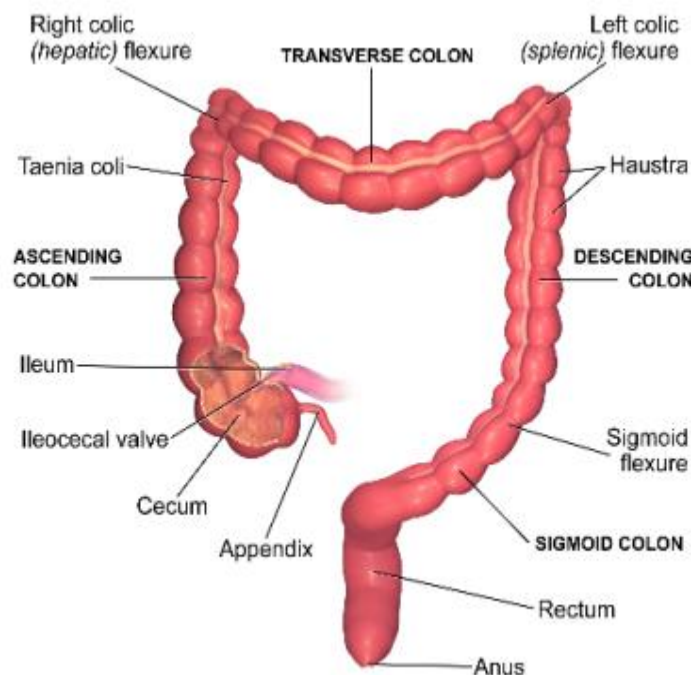


Figure 3 - Illustration du côlon humain [9]

En effet, selon sa localisation, le cancer colorectal peut prendre différentes formes.

Cette maladie se développe lentement, sur une période d'environ 10 à 20 ans. Dans la plupart des cas, ce sont des tumeurs bénignes qui apparaissent dans le côlon, qu'on appelle polypes (figure 4), très fréquentes, et qui pour certaines de ces polypes, représentent le stade précoce du cancer colorectal. [9]

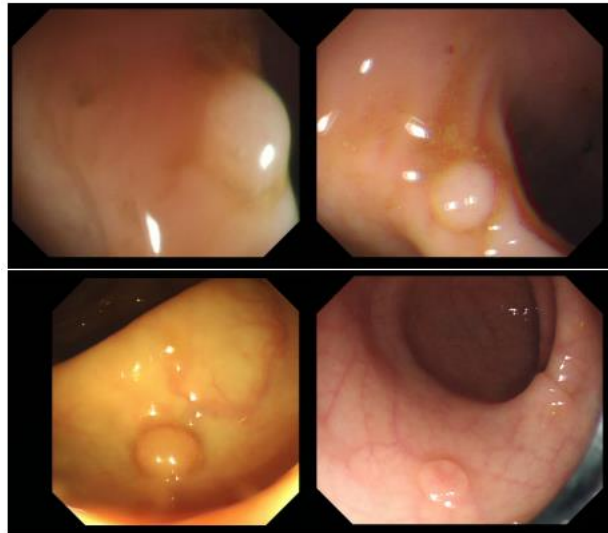


Figure 4 - Illustration concernant différentes formes de polypes [9]

Ces polypes peuvent être de différents types :

- ✓ Pédiculés : polype rattachée à la paroi du côlon par un pied, comme un champignon (figure 5-a)
- ✓ Sessiles : polype prenant la forme du sommet d'une colline posée sur la paroi (figure 5-b)
- ✓ Plans : polype presque imperceptible en endoscopie standard (figures 5-c, 5-d, 5-e, 1.5-f)

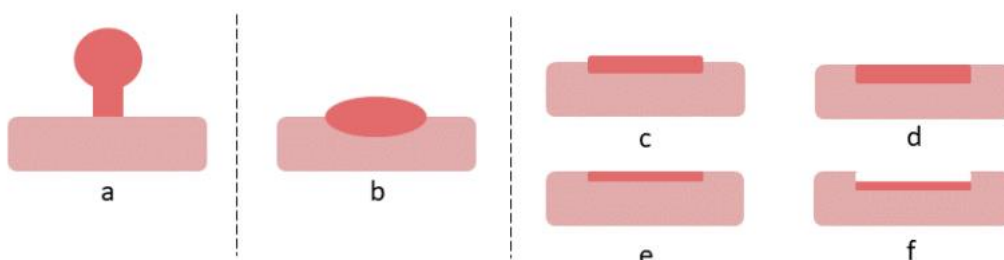


Figure 5 - Classification morphologique des polypes [9]

Ces polypes peuvent être de nature bénigne ou maligne. Lorsqu'on détecte chez un patient plus d'une dizaine de polypes, on parle alors de polypose. Du point de vue histologique (qui permet d'explorer la structure du polype et d'en trouver sa dangerosité), il existe quatre variétés de polypes [9] :

- **Les polypes adénomateux** : il s'agit de polypes qui peuvent se développer en cancer (la figure 5 montre l'évolution d'un polype adénomateux en tumeur).
- **Les polypes hyperplasiques** : il s'agit de polypes qui apparaissent naturellement et qui ne dégèrent pas nécessairement en adénomes. Cependant, un nombre élevé de ces lésions peut traduire une grande probabilité de développer des polypes adénomateux.
- **Les polypes juvéniles** : il s'agit de polypes qui sont des kystes d'origine inflammatoire.
- **Les pseudo-polypes inflammatoires** : il s'agit de polypes qui résultent d'une cicatrisation due à l'ulcération du côlon ou de la maladie de Crohn.

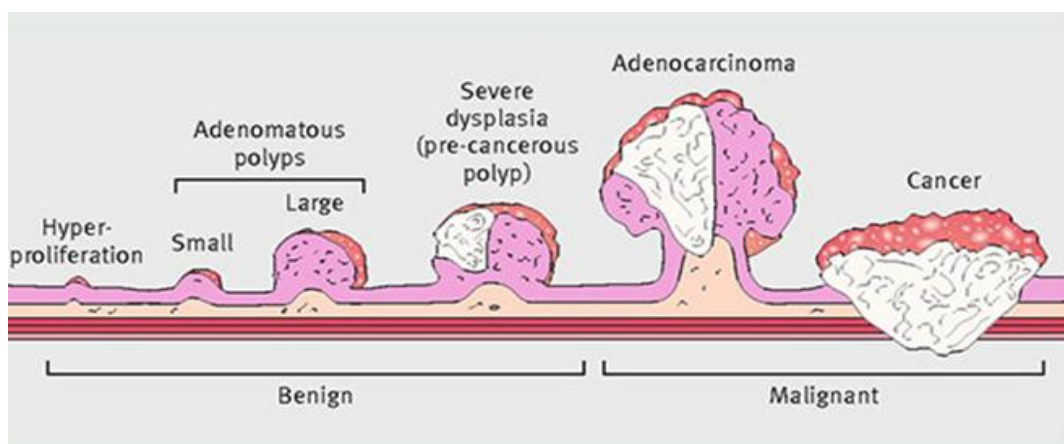


Figure 6 - illustration de l'évolution d'un polype vers une tumeur [10]

Les polypes adénomateux et l'adénocarcinome sont des tumeurs épithéliales du gros intestin et sont les néoplasmes intestinaux les plus courants et cliniquement significatifs. Le potentiel de transformation des polypes ou des adénomes en cancer augmente avec l'âge du patient. Les adénomes de plus de 1 cm, avec des motifs villositaires étendus, présentent un risque accru de développer des carcinomes. [10]

4. Détection et localisation des polypes

Dans la section précédente, plusieurs outils d'exploration du tube digestif ont été brièvement décrits. Nous nous intéressons, dans ce mémoire à la détection et à la localisation des polypes qui sont dont la présence dans le colon peut aboutir au cancer colorectal (CRC) et conduire à la mort du patient. La détection et la localisation de ces polypes revêt, ainsi, une importance vitale.

Des études montrent que l'expérience et le sens de l'observation du praticien, la qualité des images, la taille des polypes ainsi que le matériel utilisé sont parmi les facteurs influant sur la bonne détection des polypes et que très souvent un certain nombre n'est pas détecté du premier coup ou n'est pas du tout.

C'est pour pallier ces inconvénients qu'actuellement plusieurs algorithmes d'intelligence artificielle sont utilisés et aident et assistent les gastro-entérologues dans ce processus de détection-localisation des polypes.

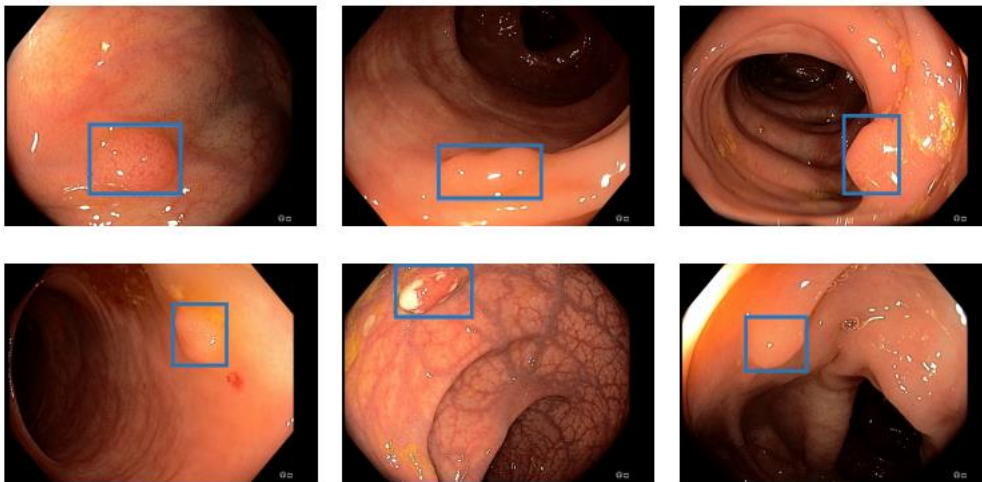


Figure 7 - illustration quelques exemples de détection des polypes. [11]

5. Apports de l'intelligence artificielle

Aujourd'hui, l'intelligence artificielle (IA) est omniprésente dans pratiquement tous les domaines de la vie sociale, économique et industrielle. En fait, intuitivement cette notion d'intelligence rejoint un vieux rêve de l'humanité qui est celui de conférer à des machines des fonctions de réflexions, de raisonnements et de comportements propres aux êtres humains. On est, ainsi, conduit à doter les machines d'algorithmes sophistiqués pour qu'elles miment le comportement humain. L'élément de base dans ce processus « d'intelligence », qui a révolutionné cette discipline, est le neurone artificiel dont le principe de fonctionnement se veut être copié de celui des neurones naturel (figure 7). Le cerveau humain est composé de milliards de neurones dont l'agencement est organisé selon un ordre bien déterminé et qui nous permet

de réfléchir, de raisonner et de résoudre plusieurs problèmes. Une autre dimension dans cette notion d'intelligence concerne l'apprentissage. Pour ce deuxième point, les machines elles-mêmes sont dotées d'algorithmes d'apprentissage automatique (Machine Learning : ML).

Le domaine de la médecine est l'un des secteurs où l'application des techniques d'IA est très recherchée. Le domaine de la gastro-entérologie bénéficie, particulièrement, des applications de l'IA pour la détection, la localisation et aussi pour l'identification pathologique. Le dépistage du cancer est un sujet brûlant et pour lequel l'IA apporte une grande aide dans l'assistance des praticiens.

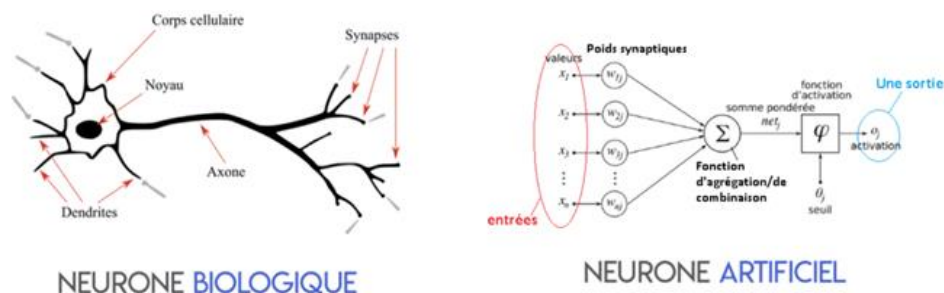
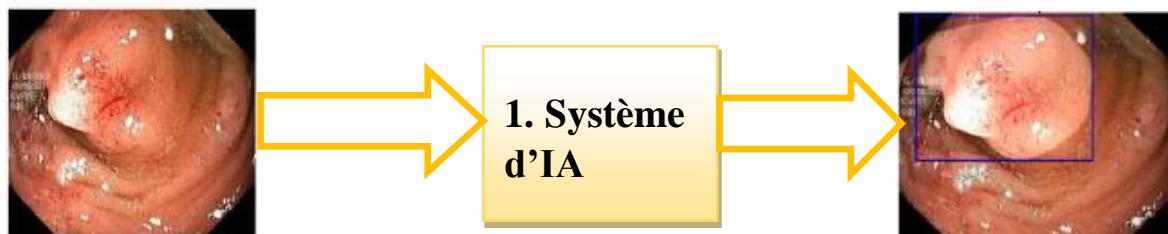


Figure 8 - comparaison entre les neurones biologiques et le neurone artificiel [12]

6. Notion de traitement des images côloscopiques



*Image numérique obtenue
Par coloscopie*

*Résultats (détection et localisation
d'éventuels polypes)*

Figure 9 - Système d'IA pour la détection et la localisation de polypes

La figure 9 montre de façon très schématique le système de détection, localisation et identification des polypes. Il comporte trois niveaux : L'image numérisée qui est obtenue grâce à la caméra embarquée dans l'outil exploratoire utilisé (fibroscope, capsule, ...). Le système d'intelligence artificielle dont le but est d'analyser l'image numérique, d'en extraire des caractéristiques et d'en déduire des résultats (exemple : présence de polypes, localisation, taille, ...). Une image numérique peut être assimilée à une matrice de points. Chaque point est dénommé pixel.

Dans le système RVB, les trois couleurs de base R (Rouge), V (Vert) et B (Bleu) sont combinés

pour former la couleur d'un pixel.

Il est aisé de constater ainsi que la taille d'une image peut être très grande et de l'ordre de millions de pixels. Chaque pixel peut, lui aussi, être codé sur plusieurs octets. Il devient alors clair que le stockage, l'exploration, le traitement et le transfert d'une image sont très consommateurs en espace de stockage et en capacité de traitement.

Pour pallier ces inconvénients, plusieurs algorithmes de compression et de segmentation d'image sont utilisés. La compression d'image permet d'obtenir des images dont les tailles sont nettement plus petites que les images originales qui peuvent être avec ou sans perte d'information (RLE, JPEG, ...), tandis que dans la segmentation d'images seules les éléments pertinents de l'image sont retenus (exemple contour, région, ...).

- Les méthodes par partitionnement : les algorithmes des k-means (k-moyennes)
- Les méthodes de regroupement hiérarchique : classification ascendante hiérarchique (CAH)

Dans notre travail, nous avons adopté des solutions qui s'appuient sur l'apprentissage supervisé.

7. Systèmes de coloscopie assistés par l'IA

7.1. Détection assistée par ordinateur (CAD) des polypes

Les algorithmes d'IA peuvent analyser les images de coloscopie et aider à la détection des polypes. Ces systèmes utilisent des techniques de traitement d'image, de reconnaissance des formes et des algorithmes d'apprentissage automatique pour identifier les polypes potentiels. Les systèmes CAD peuvent contribuer à réduire le nombre de polypes non détectés et à améliorer le taux de détection globales.

7.2. Classification des polypes

L'IA peut aider à classer les polypes détectés pendant la coloscopie. En analysant les différentes caractéristiques des polypes, telles que la forme, la texture et la couleur. Les algorithmes d'IA peuvent fournir des informations supplémentaires pour faciliter le diagnostic.

Cette classification peut aider à différencier les lésions bénignes des lésions potentiellement malignes, aidant ainsi les cliniciens à prendre des décisions éclairées.

7.3. Aide à la navigation

L'IA peut fournir une aide à la navigation en temps réel pendant les procédures de coloscopie. En analysant le flux vidéo endoscopique, les algorithmes d'IA peuvent aider à guider l'endoscopiste dans le côlon, garantissant un examen complet et réduisant le risque de polype manqué. Ces systèmes peuvent inclure des fonctions telles que la cartographie automatisée du côlon, le suivi de la position de l'endoscope et la suggestion d'itinéraires optimaux.

8. Recherche et études cliniques

- **Byrne et al. (2019)** : Cette étude évalue les performances d'un système d'IA pour la détection des polypes pendant la coloscopie. Les chercheurs ont comparé les performances d'endoscopistes experts avec celles du système d'IA dans la détection des polypes, et les résultats ont montré que le système d'IA améliorait les taux de détection des polypes. [13]
- **Wang et al. (2020)** : Dans cette étude, les chercheurs ont développé un système assisté par l'IA pour la classification des polypes lors d'une coloscopie. Ils ont entraîné un algorithme d'apprentissage profond à l'aide d'un vaste ensemble de données d'images de coloscopie et ont évalué ses performances en matière de distinction entre différents types de polypes. [14]
- **Chen et al. (2021)** : Cette étude a présenté un algorithme de détection automatique des polypes basé sur la méthode SSD utilisant un modèle VGG-16. Ils ont comparé ce modèle aux annotations réelles et aux résultats d'un Mask R-CNN. L'étude a utilisé un ensemble de 4900 images. Les résultats ont montré que le modèle SSD avait une précision moyenne (mAP) de 96 %, surpassant la détection manuelle et le Mask R-CNN. [15]
- **Wan et al. (2021)** : Ce travail présente une approche combinant le modèle YOLOv5 avec des mécanismes d'auto-attention pour détecter les polypes. Ils ont utilisé l'ensemble de données Kvasir-SEG et ont également collecté des images d'un hôpital local pour construire leur propre ensemble de données appelé WCY. Le modèle a été entraîné et évalué sur ces ensembles de données. Les résultats ont montré une précision de 92 % pour l'ensemble de données Kvasir, ainsi qu'une précision de 91 % pour l'ensemble de données WCY. [16]

- **Quan et al. (2022)** : ont développé un système d'aide à la détection des polypes, appelé EndoVigilant, basé sur une architecture de détection en un seul coup. Pour entraîner ce système, ils ont utilisé 83 000 images de coloscopie, comprenant des polypes de différentes tailles, morphologies et difficultés de détection, annotées et examinées par une équipe spécialisée. Pour valider le système EndoVigilant, 21 454 images de côlon issues d'un ensemble de données externe ont été utilisées. Le système a obtenu une sensibilité de 0,90, une spécificité de 0,97. [17]

9. Les défis et les limites associés à l'intégration des technologies de l'IA dans la coloscopie

➤ **Qualité des données**

L'efficacité des algorithmes d'IA en coloscopie dépend fortement de la qualité des données d'entraînement. Il peut être difficile d'obtenir des ensembles de données importants, diversifiés et correctement annotés. Des problèmes tels que la variabilité de la qualité de l'image, les variations de l'apparence des lésions, la taille du data set et la limitation de l'accès aux données annotées peuvent affecter les performances des modèles d'IA [18].

➤ **Considérations réglementaires**

L'intégration des technologies d'IA dans des procédures médicales telles que la coloscopie soulève des problèmes de réglementation. Le respect des réglementations et des normes, telles que celles relatives à la confidentialité des données, au consentement des patients et à l'approbation des dispositifs médicaux, peut s'avérer complexe. Il est essentiel de s'assurer que les systèmes d'IA répondent aux exigences réglementaires pour que leur déploiement soit sûr et éthique.

➤ **Validation des algorithmes d'IA**

Il est essentiel de valider les performances et la généralisation des algorithmes d'IA en coloscopie. Des études cliniques et des évaluations comparatives solides sont nécessaires pour démontrer l'efficacité, la fiabilité et la sécurité des systèmes de coloscopie assistés par l'IA. Des protocoles d'évaluation normalisés et des critères de référence sont nécessaires pour garantir des évaluations cohérentes et significatives [19].

➤ **Questions éthiques**

L'utilisation de l'IA en coloscopie soulève des questions éthiques. La confidentialité des patients et la sécurité des données doivent être préservées tout au long du flux de travail de l'IA, depuis la collecte et le stockage des données jusqu'à la formation et au déploiement des algorithmes. En outre, des questions telles que la partialité des algorithmes, la transparence, la responsabilité et l'impact sur les relations médecin-patient doivent être examinées avec soin [20].

➤ **Intégration clinique et flux de travail**

L'intégration des technologies d'IA dans le flux de travail clinique pose des problèmes. L'intégration transparente des systèmes d'IA à l'équipement de coloscopie, à l'acquisition d'images et aux systèmes de rapports existants peut s'avérer complexe. Il est essentiel de veiller à ce que les algorithmes d'IA fournissent des informations exploitables sans perturber l'efficacité et l'efficacité de la procédure [21].

➤ **Collaboration entre l'homme et l'IA**

Il est essentiel de parvenir à une collaboration efficace entre les systèmes d'IA et les professionnels de la santé. Les cliniciens doivent comprendre et faire confiance aux résultats des algorithmes d'IA, ce qui nécessite de pouvoir les expliquer et les interpréter. Des programmes de formation et d'éducation efficaces sont nécessaires pour améliorer les compétences et les connaissances des prestataires de soins de santé en matière d'utilisation des technologies d'IA [22].

10. Impacts potentiels de l'IA en coloscopie

➤ **Amélioration de la détection des polypes**

Les algorithmes d'IA se sont révélés prometteurs pour améliorer la détection des polypes et des lésions pendant la coloscopie. Les recherches futures visent à affiner et à optimiser ces algorithmes, afin d'accroître la sensibilité et la spécificité. Il pourrait en résulter de meilleurs taux de détection et une réduction des taux d'échec, ce qui améliorerait en fin de compte les résultats pour les patients.

➤ **Aide à la décision en temps réel**

L'IA a le potentiel de fournir une aide à la décision en temps réel pendant les procédures de coloscopie. En analysant les images et les vidéos endoscopiques en temps réel, les

algorithmes d'IA peuvent aider les cliniciens à identifier les zones suspectes, à recommander des actions appropriées et à fournir des conseils pour un diagnostic et un traitement précis.

➤ **Médecine de précision**

Avec l'intégration de l'IA, la coloscopie peut évoluer vers une approche plus personnalisée. Les algorithmes d'IA peuvent analyser les antécédents médicaux, les informations génétiques et d'autres données pertinentes d'un patient afin d'adapter les stratégies de dépistage et de surveillance en fonction des profils de risque individuels.

Cela peut conduire à des interventions plus ciblées et plus efficaces, en réduisant les procédures inutiles pour les personnes à faible risque et en assurant une détection opportune pour les patients à haut risque [23].

➤ **Coloscopie virtuelle**

La coloscopie virtuelle, également connue sous le nom de colonographie par tomographie assistée par ordinateur (CTC), est une méthode alternative non invasive pour l'examen du côlon. L'IA peut jouer un rôle dans l'amélioration de la précision de la coloscopie virtuelle en aidant à la détection et à la caractérisation des lésions à partir des images de tomodensitométrie. Les recherches en cours se concentrent sur le développement d'algorithmes d'IA pour des procédures de coloscopie virtuelle efficaces et fiables.

➤ **Intégration aux dispositifs endoscopiques**

Les technologies de l'IA peuvent être intégrées aux dispositifs endoscopiques afin d'améliorer leurs capacités. Par exemple, les systèmes de navigation guidés par l'IA peuvent aider à guider les endoscopes dans le côlon, améliorant ainsi la maniabilité et réduisant la durée de la procédure. En outre, les systèmes robotiques alimentés par l'IA ont le potentiel d'automatiser certaines tâches pendant la coloscopie, ce qui accroît l'efficacité et réduit les erreurs humaines [24].

➤ **Amélioration des taux de détection**

Les algorithmes d'IA peuvent contribuer à améliorer les taux de détection des polypes et des lésions, réduisant ainsi le risque de diagnostics manqués et améliorant la détection précoce du cancer colorectal.

➤ **Réduction de la durée de la procédure**

Grâce aux systèmes guidés par l'IA, les procédures de coloscopie peuvent devenir plus efficaces et rationalisées. Cela peut conduire à des temps de procédure plus courts, à un

meilleur confort pour le patient et à une augmentation du débit dans les établissements de soins de santé.

➤ **Réduction des coûts**

La coloscopie assistée par l'IA a le potentiel de réduire les coûts des soins de santé en minimisant le besoin de procédures inutiles, en optimisant l'allocation des ressources et en améliorant les résultats pour les patients. En détectant et en traitant les anomalies à des stades plus précoces, l'IA peut potentiellement réduire le fardeau économique associé au traitement des maladies à un stade avancé [25].

11. Conclusion

En conclusion, l'intégration de l'intelligence artificielle dans la coloscopie offre de nombreuses perspectives pour améliorer la détection précoce et la classification des lésions précancéreuses du côlon. Les systèmes de coloscopie assistés par l'IA peuvent aider les médecins à repérer les lésions plus facilement, à les caractériser avec plus de précision et à guider les biopsies et les résections. Cependant, malgré les avancées technologiques et les résultats prometteurs des études cliniques, il reste des défis à relever pour une intégration optimale de l'IA dans la pratique clinique. La formation des médecins, la standardisation des protocoles et la validation des algorithmes sont autant d'enjeux à prendre en compte pour garantir l'efficacité et la sécurité de ces nouvelles technologies.

Chapitre 02: L'apprentissage profond (Deep-Learning: DL)

1. Introduction

Dans ce deuxième chapitre, nous allons explorer les fondements de l'intelligence artificielle, en nous concentrant sur l'apprentissage automatique et l'apprentissage profond, ainsi que sur la vision par ordinateur et les modèles de détection d'objets par CNN. Nous allons également décrire les différentes méthodes de détection d'objets, en mettant l'accent sur la méthode YOLO, qui est connue pour sa rapidité et sa précision. En comprenant ces concepts clés, nous serons mieux équipés pour explorer les applications de l'intelligence artificielle dans notre vie quotidienne.

2. L'intelligence artificielle

L'intelligence artificielle est un domaine dont les retombées occupent de plus en plus de place dans un très grand nombre d'applications. Il s'agit d'équiper les machines avec des algorithmes et du matériel leur permettant de reproduire le comportement humain [26]. L'objectif étendu de l'IA est de créer des machines capables d'effectuer des tâches complexes de manière autonome, en s'adaptant à leur environnement et en prenant des décisions intelligentes.

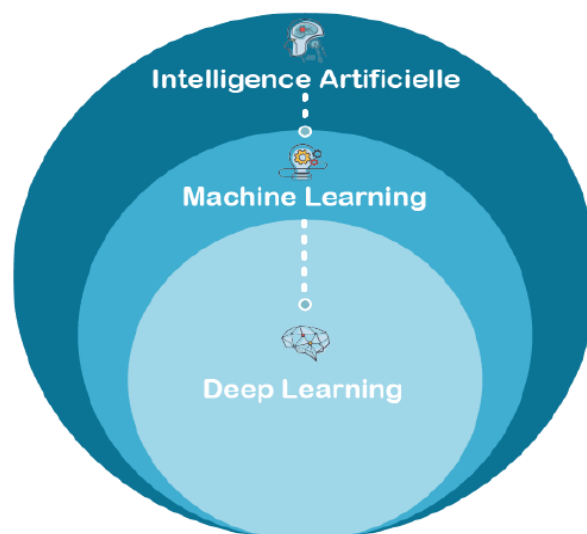


Figure 10 - Hiérarchie des domaines de l'intelligence artificielle [27]

3. L'Apprentissage automatique

L'apprentissage automatique, également connu sous le nom de machine learning (ML), est un domaine d'étude de l'intelligence artificielle qui repose sur des approches statistiques pour permettre aux ordinateurs d'apprendre à partir de données, sans être explicitement programmés pour chaque tâche.

Il se déroule généralement en deux phases, La première phase consiste à créer un modèle à partir des données, appelée phase d'apprentissage. Une fois le modèle établi, la deuxième phase, appelée mise en production, permet d'utiliser le modèle pour obtenir des résultats en traitant de nouvelles données pour les prédictions [28].

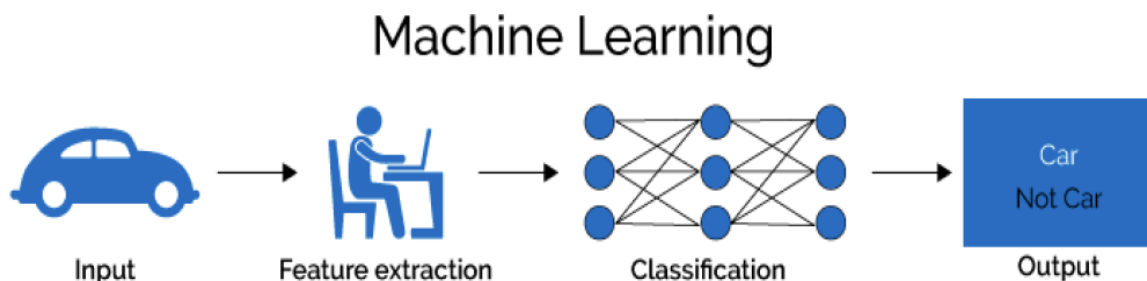


Figure 11- Architecture d'apprentissage automatique [29]

Trois types couramment utilisés d'apprentissage machine sont l'apprentissage par renforcement, l'apprentissage supervisé et l'apprentissage non supervisé.

3.1. L'apprentissage supervisé

Les algorithmes d'apprentissage automatique qui mettent en œuvre un apprentissage par paire des données d'entrée et de sortie (inputs/outputs) sont appelés « supervisés », car il y a une intervention et une supervision humaines.

L'apprentissage supervisé est actuellement la méthode privilégiée et généralement la plus performante pour la détection d'objets dans les images. Compte tenu des résultats favorables obtenus, ces dernières années, l'apprentissage supervisé a connu des avancées significatives et remarquables.

Parmi les algorithmes d'apprentissage automatique supervisés les plus utilisés, on trouve : Les arbres de décision, les k plus proches voisins, les machines à vecteurs, la régression logistique, la classification naïve bayésienne et les réseaux de neurones profonds. [30]

3.2. L'apprentissage non supervisé

La deuxième catégorie d'algorithmes d'apprentissage automatique est appelée "apprentissage non supervisé". Dans ce cas, les données ne sont pas préalablement étiquetées, et l'algorithme est laissé libre de tirer ses propres conclusions.

Ce type d'apprentissage revêt une grande importance, car il est plus courant dans le fonctionnement du cerveau humain que l'apprentissage supervisé.

Les algorithmes d'apprentissage non supervisé sont particulièrement utilisés dans les problèmes de regroupement (clustering), où l'objectif est de comprendre et de représenter les relations entre les objets d'une collection donnée. [30]

3.3. L'apprentissage par renforcement

L'apprentissage par renforcement est une technique qui permet à un agent d'apprendre à agir dans son environnement en maximisant une récompense. L'agent perçoit son environnement, choisit une action et reçoit une récompense en fonction de cette action. L'objectif de l'apprentissage par renforcement est de trouver la meilleure action à effectuer dans chaque situation (état) afin de maximiser les récompenses.

Cette méthode d'apprentissage est basée sur des essais et des erreurs, et elle ne nécessite pas de supervision externe pour indiquer l'action optimale. Elle utilise plutôt la récompense comme évaluation de la qualité de l'action choisie. [31]

4. L'apprentissage profond

L'apprentissage profond ou deep learning (DL) est un type particulier d'apprentissage automatique qui se concentre sur l'utilisation d'algorithmes appelés réseaux de neurones artificiels, s'inspirant ainsi de la structure et du fonctionnement du cerveau humain. [32]

Dans le contexte de l'apprentissage profond, plusieurs couches sont utilisées pour extraire des caractéristiques de haut niveau et importantes à partir des données d'entraînement. Par exemple, dans le traitement d'images, les couches inférieures peuvent détecter les contours, tandis que les couches supérieures peuvent reconnaître des concepts plus complexes tels que les chiffres, les lettres ou les visages. [33]

La principale utilité du deep learning réside dans sa capacité à traiter de grandes quantités de données afin d'améliorer les performances d'un algorithme et de rendre un système de plus en plus intelligent et autonome.

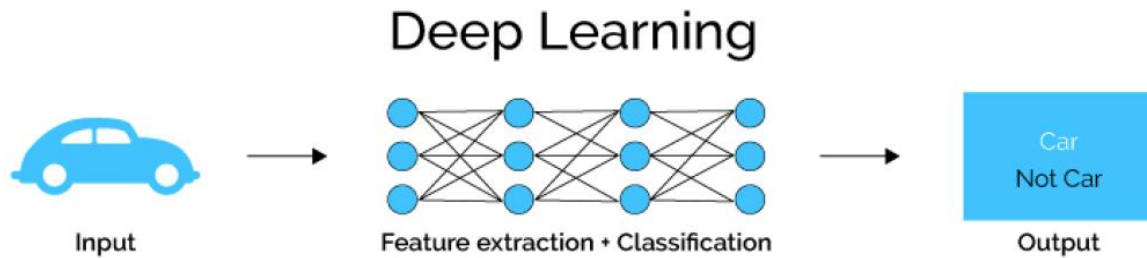


Figure 12 - Architecture d'apprentissage profond [34]

L'architecture, illustrée Figure 12, comprend plusieurs couches qui sont responsables de l'extraction des caractéristiques et de la classification, permettant au réseau d'apprendre et d'extraire des caractéristiques significatives à partir de l'ensemble de données d'entrée.

4.1. Les réseaux de neurones profonds

Les réseaux de neurones profonds, également connus sous le nom de Deep Neural Networks (DNN), Ces réseaux sont semblables aux réseaux multi couches (MLP : Multi Layer Network) en comportant plusieurs couches cachées. Cette augmentation du nombre de couches, permet au réseau de neurones de détecter les légères variations du modèle d'apprentissage et favorise le sur-apprentissage ou sur-ajustement (« *overfitting* »).

Les réseaux de neurones profonds sont largement utilisés dans des domaines tels que la vision par ordinateur, le traitement du langage naturel et d'autres tâches de perception et de compréhension des données complexes. Ils ont démontré leur efficacité dans de nombreux problèmes, mais nécessitent souvent une quantité importante de données d'apprentissage et une puissance de calcul suffisante pour leur entraînement. [35]

4.2. Les réseaux de neurones convolutifs

Les réseaux de neurones convolutifs sont une catégorie de réseaux neuronaux qui se concentrent principalement sur l'analyse visuelle des images, notamment la reconnaissance et la classification d'images. Lorsque nous observons une image, notre cerveau humain traite simultanément et de manière complexe une multitude d'informations pour en extraire du sens. Cette capacité des CNN tente de mimer le fonctionnement du cerveau, qui sont conçus pour imiter ce processus en utilisant des techniques de convolution pour détecter des motifs et des caractéristiques spécifiques dans les images.

En comparaison avec d'autres types de réseaux neuronaux, les CNN tirent profit de leur structure spéciale, où la connectivité locale réduite entre les couches permet une réduction du

nombre de paramètres, ce qui accélère le calcul. De plus, les poids partagés, tels que les filtres de noyau, permettent de capturer les propriétés locales du signal, ce qui rend les CNN particulièrement adaptés à la tâche de l'analyse visuelle des images. [36]

4.2.1. Architecture d'un réseau de neurones convolutive

Les CNN sont actuellement des modèles utilisés en classification d'images qui donnent des résultats très performants. L'utilisateur fournit en entrée une image sous la forme d'une matrice de pixels, en deux dimensions lorsqu'il s'agit d'une image en niveau de gris ou en trois dimensions lorsque c'est une image couleur, la troisième dimension permettant de représenter les trois couleurs de base rouge, vert et bleu (RVB). L'architecture du réseau de neurones convolutif dispose en amont d'une partie convolutive et comporte par conséquent deux parties bien distinctes :

- **Une partie convolutive** : dont l'objectif final est d'extraire des caractéristiques propres à chaque image en les compressant de façon à réduire leur taille initiale. L'image lue par le réseau traverse une succession de filtres, créant par la même occasion de nouvelles images appelées cartes de convolutions. Enfin, les cartes de convolutions obtenues sont concaténées dans un vecteur de caractéristiques appelé code CNN.
- **Une partie classification** : Le code CNN obtenu en sortie de la partie convolutive est fourni en entrée dans une deuxième partie, constituée de couches entièrement connectées appelées perceptron multicouche (MLP pour Multi Layers Perceptron). Le rôle de cette partie est de combiner les caractéristiques du code CNN afin de classer l'image. [37]

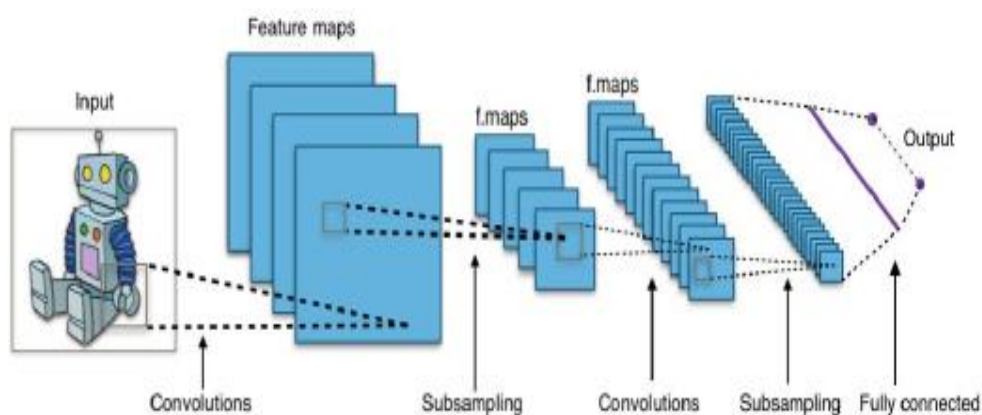


Figure 13 - Architecture d'un réseau de neurones convolutive [38]

Une architecture CNN est formée par un empilement de couches de traitement indépendantes :

➤ **La couche convolution**

La couche de convolution, elle permet d'extraire des caractéristiques significatives de l'image en effectuant une opération de convolution.

Dans cette couche, une portion de l'image est connectée à la couche de convolution. Cette portion de l'image, appelée champ récepteur, a la même taille que le filtre utilisé dans la convolution. L'opération de convolution consiste à calculer le produit scalaire entre le champ récepteur et le filtre. Le résultat de cette opération est un entier unique qui constitue une valeur dans le volume de sortie.

Ensuite, le filtre glisse sur le champ récepteur suivant de l'image d'entrée en utilisant un décalage. On répète alors la même opération de convolution pour cette nouvelle position du filtre. Ce processus est répété jusqu'à ce que l'ensemble de l'image ait été parcouru. [39]

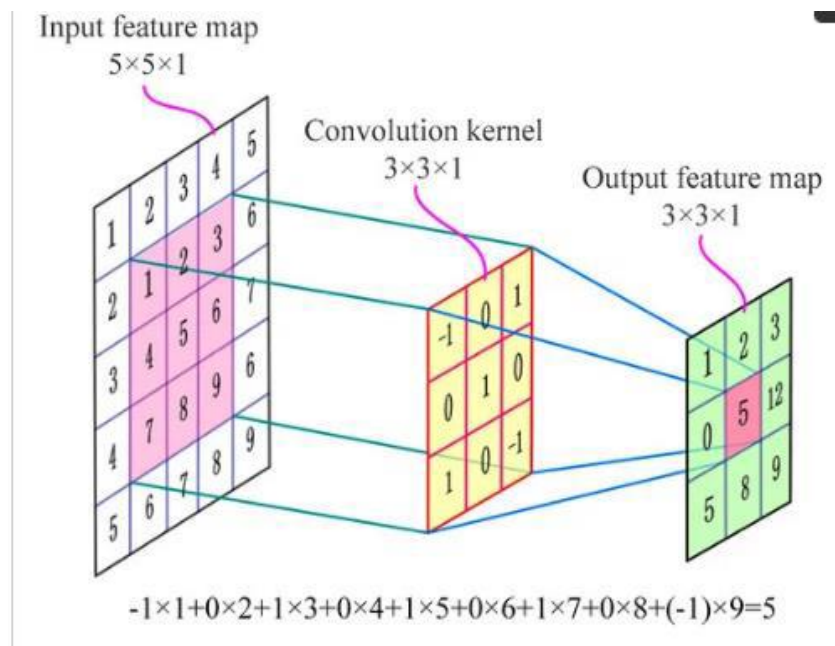


Figure 14 - La couche de convolution [25]

➤ **Couche de mise en commun (Pooling layers)**

La couche de mise en commun est une opération de réduction de taille, généralement utilisée entre deux couches de convolution.

Cette couche permet de réduire progressivement la taille de l'image en préservant les informations les plus importantes.

Elle utilise généralement le Max Pooling, qui consiste à sélectionner le pixel ayant la valeur maximale dans chaque groupe (par exemple de 4 pixels), ou l'AVG Pooling, qui calcule la moyenne des pixels dans le groupe. [38]

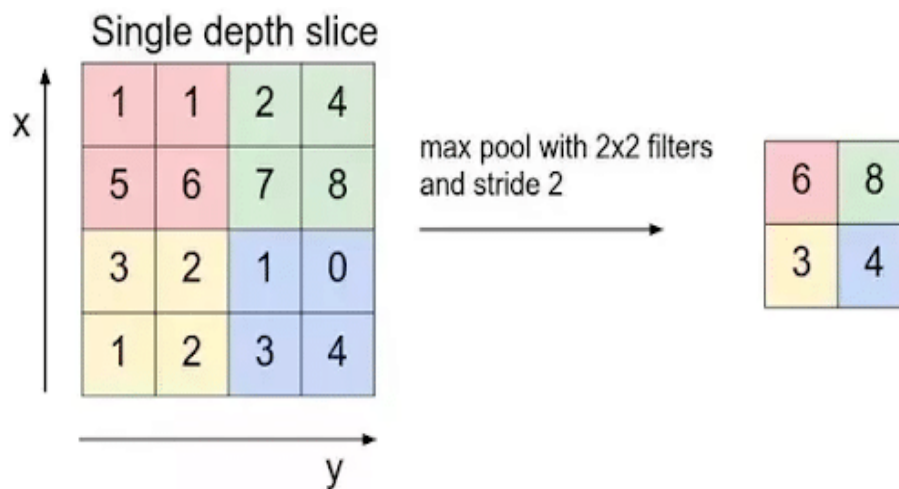


Figure 15- le Fonctionnement de la couche Pooling [27]

Avec la couche pooling, la quantité de paramètres et de calcul dans le réseau est réduite, ce qui aide à contrôler le sur-apprentissage.

➤ **Couches de correction (ReLU)**

L'efficacité du traitement peut être améliorée en intercalant entre les couches de traitement une couche dite de correction ReLU (Rectified Linear Units) qui va opérer une fonction mathématique (fonction d'activation) sur les signaux de sortie. Exemple : la fonction ReLU $F(x)=\max(0, x)$ force les neurones à retourner des valeurs positives (toutes les valeurs négatives sont remplacées par 0). [28]

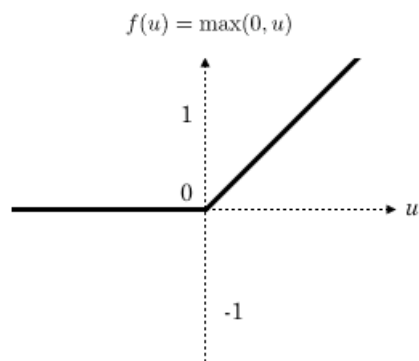


Figure 16 - Fonction d'activation ReLU [29]

➤ **Couche entièrement connectée (Fully Connected)**

Les couches entièrement connectées fonctionnent comme un réseau de neurones traditionnel et contiennent environ 90 % des paramètres d'un CNN.

Les couches entièrement connectées dans un réseau neuronal sont responsables du raisonnement de haut niveau. Les neurones de ces couches sont connectés à tous les neurones de la couche précédente. Pour calculer leurs activations, une multiplication matricielle est effectuée suivie d'un décalage de polarisation. [30]

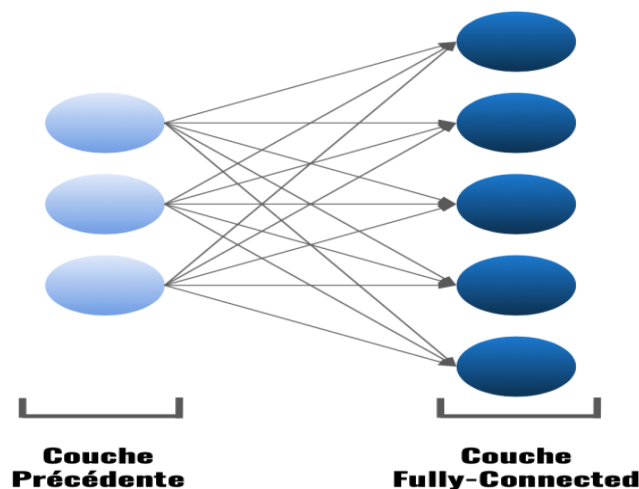


Figure 17- Couche entièrement connectée [27]

➤ **La couche de perte (LOSS)**

C'est à ce niveau qu'est spécifié l'écart entre le signal prévu et le signal réel. C'est normalement la dernière couche dans le réseau. On distingue plusieurs fonctions de perte adaptées à des tâches différentes qui peuvent y être utilisées. Par exemple, la fonction « Softmax » permet de calculer la distribution de probabilités sur les classes de sortie.

5. Vision par ordinateur

La vision par ordinateur (CV) est définie comme un domaine d'étude interdisciplinaire visant à aider les ordinateurs à visualiser, analyser et comprendre le contenu des images et des vidéos. Ce problème semble simple car les personnes peuvent le résoudre facilement et sans effort. Cependant, il reste toujours un problème partiellement résolu en raison de la complexité de la perception visuelle. [31]

La vision par ordinateur est donc un domaine d'étude axé sur l'objectif d'aider les ordinateurs à voir. On peut la classer de manière générale comme une sous-branche de l'intelligence artificielle et de l'apprentissage automatique. [31]

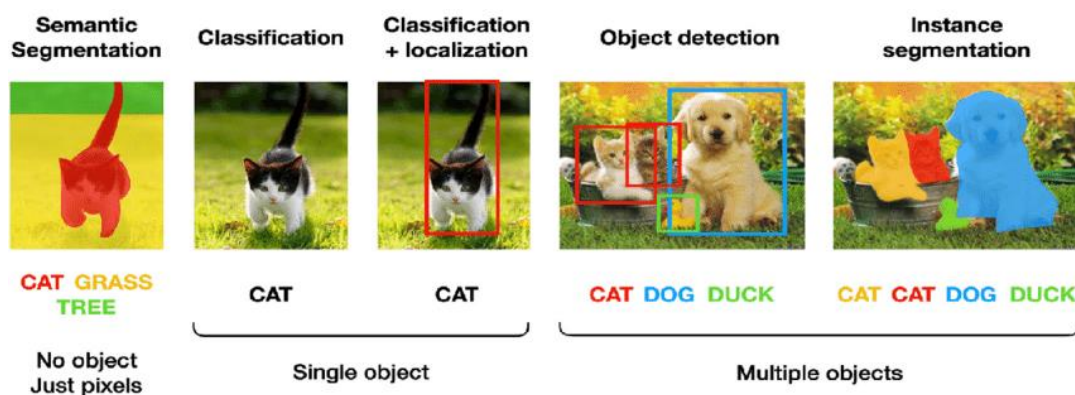


Figure 18 - Tâches de vision par ordinateur [32]

La classification d'images prédit la catégorie d'un objet dans une image, tandis que la localisation d'objets identifie précisément leur position. La détection d'objets combine ces deux tâches en localisant et classifiant les objets simultanément. Ci-dessous une brève explication en termes d'entrée et de sortie des tâches de vision par ordinateur [31].

- **Classification d'image :** Prédire le type ou la classe d'un objet dans une image. L'entrée est une image avec un seul objet. La sortie est une étiquette de classe spécifiant le type de classification de l'image.
- **Localisation d'objet :** Localiser la présence d'objets dans une image et indiquer leur emplacement à l'aide d'une boîte englobante. L'entrée est une image avec un ou plusieurs objets. La sortie est une ou plusieurs boîtes englobantes spécifiant l'emplacement des objets dans l'image.
- **Détection d'objets :** Localiser la présence d'objets à l'aide d'une boîte englobante et indiquer les types ou les classes des objets identifiés dans une image. L'entrée est une image avec un ou plusieurs objets. La sortie est une ou plusieurs boîtes

englobantes avec une étiquette de classe attribuée à chaque boîte englobante pour spécifier l'emplacement et la classe des objets détectés.

- **Segmentation d'objet** : Mettre en évidence les pixels spécifiques appartenant aux objets localisés dans une image. L'entrée est une image avec des objets. La sortie consiste à mettre en évidence les boîtes englobantes pour les objets dans l'image.

Les architectures de réseaux de neurones convolutifs (CNN) ont la capacité d'apprendre des caractéristiques complexes. En ce qui concerne la détection d'objets, on distingue deux types de modèles basés sur les CNN, Le premier type la détection en deux coup est basé sur la proposition de région et comprend des modèles tels que RCNN, SPP-NET, FRCNN, Faster RCNN et le deuxième type la détection à un coût est basée sur la régression et comprend MultiBox, AttentionNet, G-CNN, YOLO, SSD...etc.

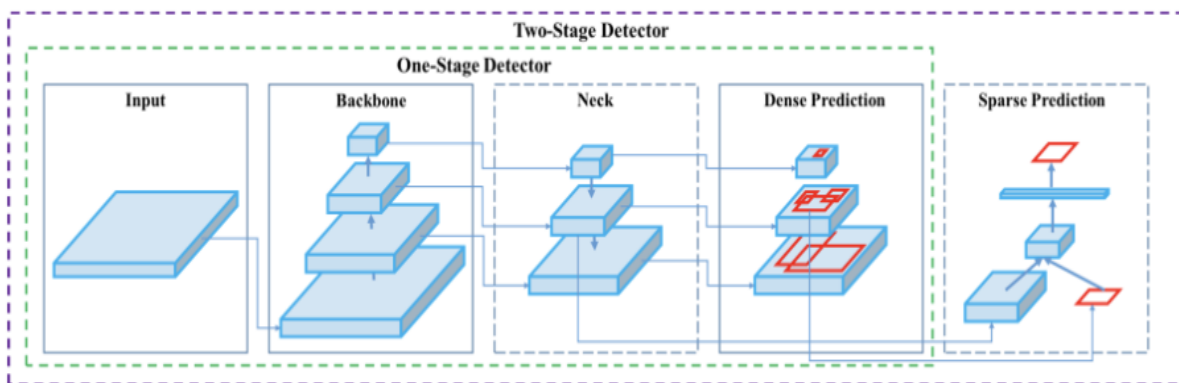


Figure 19 - Structure typique du détecteur d'objets [33]

5.1. La détection en deux étapes

Comme son nom l'indique, cette méthode comporte deux étapes. La première est la proposition de régions, puis, dans la deuxième étape, la classification de ces régions et le raffinement de la prédiction de l'emplacement ont lieu.

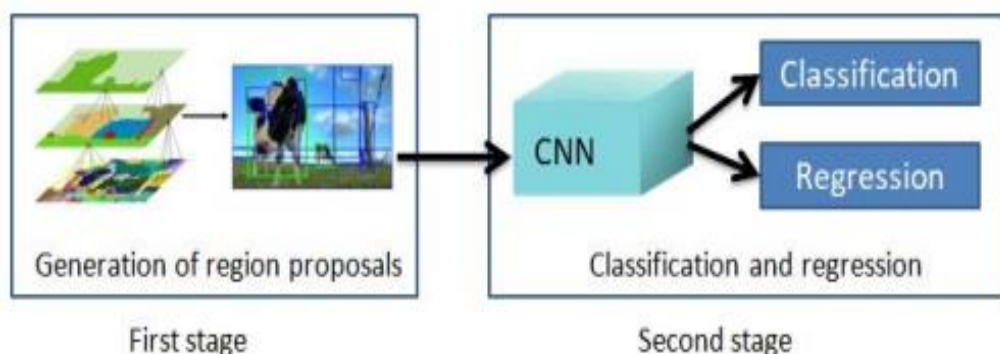


Figure 20 - Approche en deux étapes.

5.1.1. R-CNN

R-CNN (Region proposal with convolutional neural network) est le premier modèle de détection d'objets en deux étapes basées sur CNN. L'architecture de R-CNN comprend trois modules distincts, comme illustré dans la figure 21[34].

- ✓ **Génération de propositions régionales** : Ce module utilise l'algorithme de recherche sélective pour générer environ 2000 propositions de régions indépendantes de classe pour chaque image d'entrée.
- ✓ **Extraction de caractéristiques** : Ce module utilise un CNN avec cinq couches de convolution et deux couches entièrement connectées pour extraire des vecteurs de caractéristiques de longueur fixe à partir de chaque proposition de région.
- ✓ **Classification et localisation** : Ce module utilise une machine à vecteurs de support linéaire spécifique (SVM) à la catégorie pour classer chaque proposition de région.

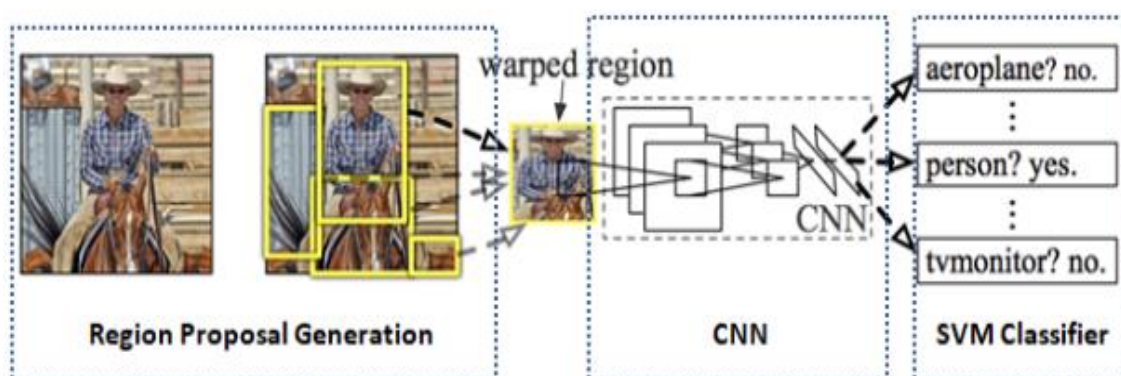


Figure 21- Architecture de modèle R-CNN [24]

5.1.2. Fast R-CNN

R-CNN et SPP-net rencontraient des problèmes communs, notamment un entraînement en plusieurs étapes, un coût élevé en termes d'espace et de temps, ainsi qu'une détection d'objets lente. Pour remédier à ces problèmes, Ross Girshick a développé Fast R-CNN, une amélioration de R-CNN qui vise à résoudre ces limitations.

Fast R-CNN permet d'apprendre simultanément la classification des propositions de régions et la correction de leurs positions spatiales. Il peut entraîner des réseaux de détection profonds tels que VGG-16 avec efficacité, offrant des gains

Significatifs de vitesse par rapport à R-CNN et SPP-net. [34]

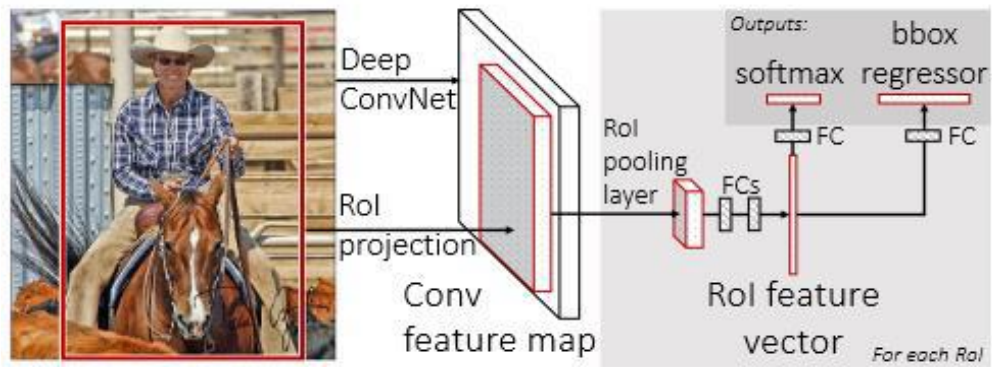


Figure 22- architecture de model Fast R-CNN [25]

La figure 22 illustre l'architecture de Fast R-CNN, Au lieu d'utiliser un CNN pour chaque proposition de région générée comme dans R-CNN, Fast R-CNN prend l'image entière et un ensemble de propositions de régions d'objets comme entrée. À partir de la carte des caractéristiques produites par le CNN, les régions d'intérêt (RoI) sont identifiées à l'aide de la méthode de recherche sélective.

Ensuite, une couche de pooling RoI est utilisée pour remodeler les régions d'intérêt en un vecteur de caractéristiques de longueur fixe. Les vecteurs de caractéristiques sont ensuite alimentés dans des couches entièrement connectées qui génèrent deux sorties : une pour la classification des objets et une autre pour la régression des boîtes englobantes.

5.1.3. Faster R-CNN

R-CNN, SPP-net et Fast R-CNN dépendent de l'algorithme de proposition de région pour la détection d'objets. Les trois modèles ont rencontré des problèmes liés au temps de calcul des propositions de région, ce qui a un impact sur les performances globales du réseau.

Shaoqing Ren et al en 2016, ont proposé Faster R-CNN, où ils ont remplacé la méthode de proposition de région précédemment mentionnée par un réseau de proposition de région (RPN). Un RPN est un réseau entièrement convolutif (FCN) qui prend une image de taille arbitraire en entrée et génère un ensemble de propositions d'objets rectangulaires. Chaque proposition d'objet est associée à un score d'objet pour déterminer si la proposition contient un objet ou non. [34]

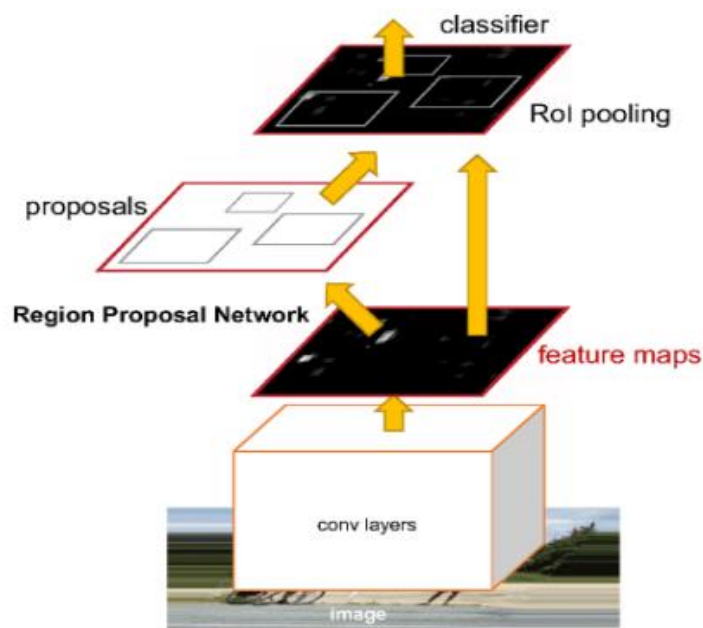


Figure 23 - Architecture de modèle Faster R-CNN [38]

De manière similaire à Fast R-CNN, l'image entière est fournie en entrée aux couches convolutives de Faster R-CNN pour produire une carte de caractéristiques convolutives. Au lieu d'utiliser un algorithme de recherche sélective sur la carte de caractéristiques pour identifier les propositions de région, un RPN est utilisé pour prédire les propositions de région. Les propositions de région détectées par fenêtres glissantes dans le RPN sont appelées ancres. Une boîte d'ancrage pertinente est sélectionnée en appliquant un seuil sur le score de "pertinence à l'objet". Les boîtes d'ancrage sélectionnées et les cartes de caractéristiques calculées par le modèle CNN initial sont alimentées à la couche de RoI pooling pour un remodelage. La sortie de la couche de RoI pooling est ensuite envoyée aux couches FC pour la classification finale et la régression des boîtes englobantes. [37]

5.2. La détection en une étape

Ce type de modèle effectue la détection et la localisation des objets en une seule passe, en utilisant des techniques de régression pour prédire directement les emplacements des objets.

5.2.1. YOLO (You Only Look Once)

YOLO (You Only Look Once), est un algorithme populaire de détection d'objets utilisé par des chercheurs du monde entier. Il a été initialement décrit en 2015 dans un article rédigé par Joseph Redmon et ses collègues [26]. L'algorithme utilise les caractéristiques de l'image dans son ensemble pour prédire les boîtes englobantes des objets. De plus, il prédit simultanément les

boîtes englobantes pour toutes les classes d'objets présentes dans l'image. Cela signifie que le réseau analyse globalement l'image entière et tous les objets qu'elle contient. La conception de YOLO permet un apprentissage de bout en bout et offre des performances en temps réel, tout en maintenant une précision moyenne élevée [49].

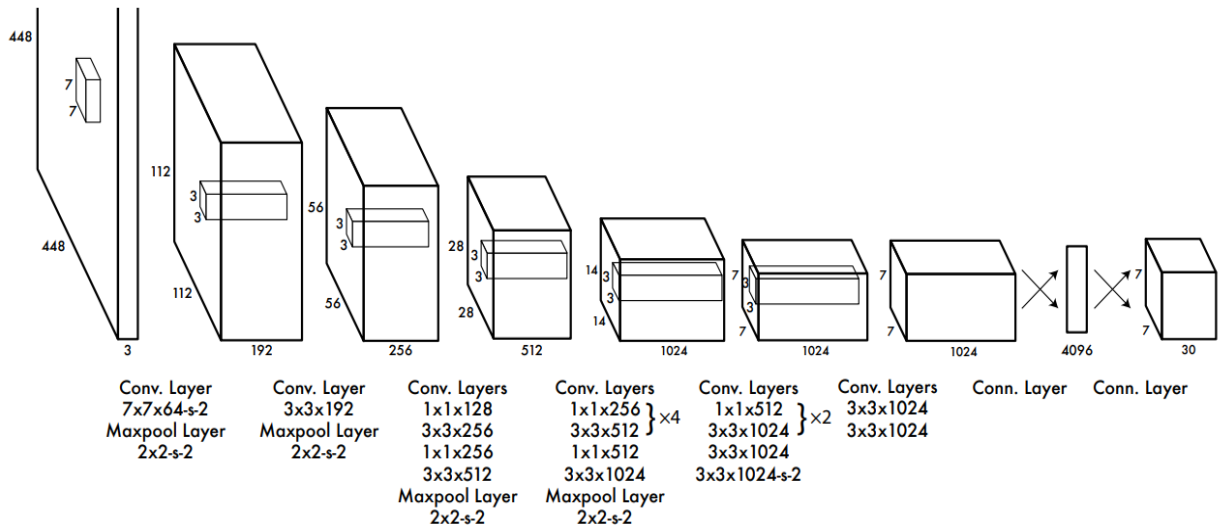


Figure 24 - Architecture de base du modèle YOLO [50]

L'architecture de ce réseau s'inspire du modèle GoogLeNet, qui est utilisé pour la classification d'images. Le réseau YOLO comprend 24 couches convolutives suivies de 2 couches entièrement connectées. Cependant, au lieu d'utiliser les modules d'initialisation présents dans GoogLeNet, YOLO utilise des couches de réduction de dimension 1×1 suivies de couches convolutives 3×3 [50].

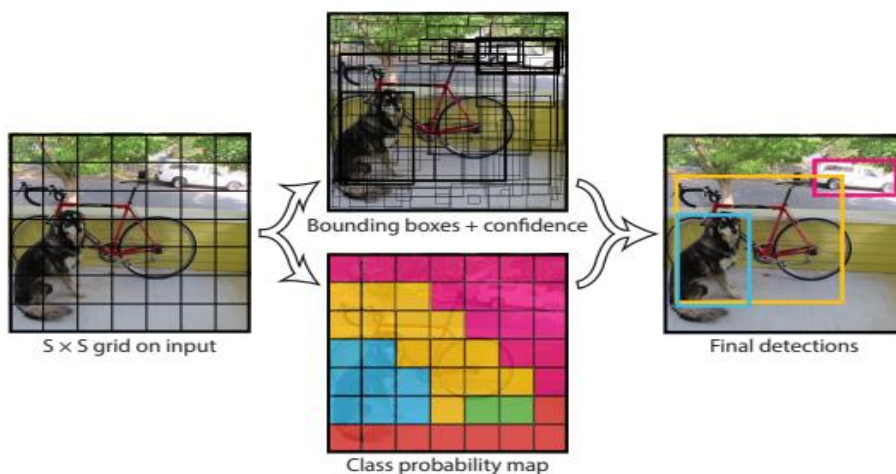


Figure 25 - Schéma du fonctionnement de YOLO [51]

Le système YOLO divise l'image en une grille de taille $S \times S$ (exemple 3×3), ce qui donne N cellules au total. Chaque cellule de la grille est chargée de détecter des objets spécifiques. Au sein de chaque cellule, plusieurs boîtes de délimitation sont prédites, accompagnées de scores de confiance pour chacune de ces boîtes. Enfin, les boîtes ayant les scores de confiance les plus élevés et les probabilités de classe les plus élevées sont considérées comme les détections finales. [52]

A. Bounding box

Une boîte englobante (bounding box) est un contour qui encadre un objet dans une image. Chaque boîte dans l'image est caractérisée par les attributs suivants :

- Largeur (b_w) box width
- Hauteur (b_h) box height
- Classe (par exemple, personne, voiture, feux de circulation...) Représente par la lettre c
- Le Centre du box (b_x, b_y)

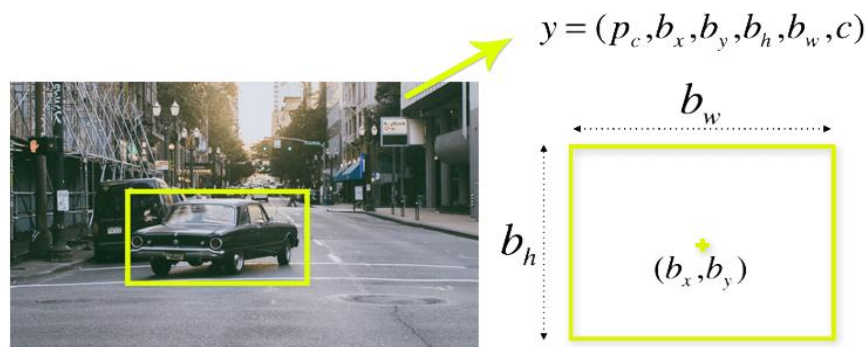


Figure 26 - boîte englobant [51]

B. Intersection over union (IoU)

L'intersection sur union (IoU) est une mesure utilisée en détection d'objets pour évaluer le degré de chevauchement entre les boîtes englobantes prédites et réelles. Dans le contexte de YOLO, chaque cellule de la grille prédit une boîte englobante et utilise l'IoU pour évaluer sa similarité avec la boîte réelle. Une IoU de 1 indique une correspondance parfaite. En utilisant l'IoU comme critère, les boîtes qui ne correspondent pas étroitement à la boîte réelle peuvent être éliminées, améliorant ainsi la précision de la détection des objets. [53]

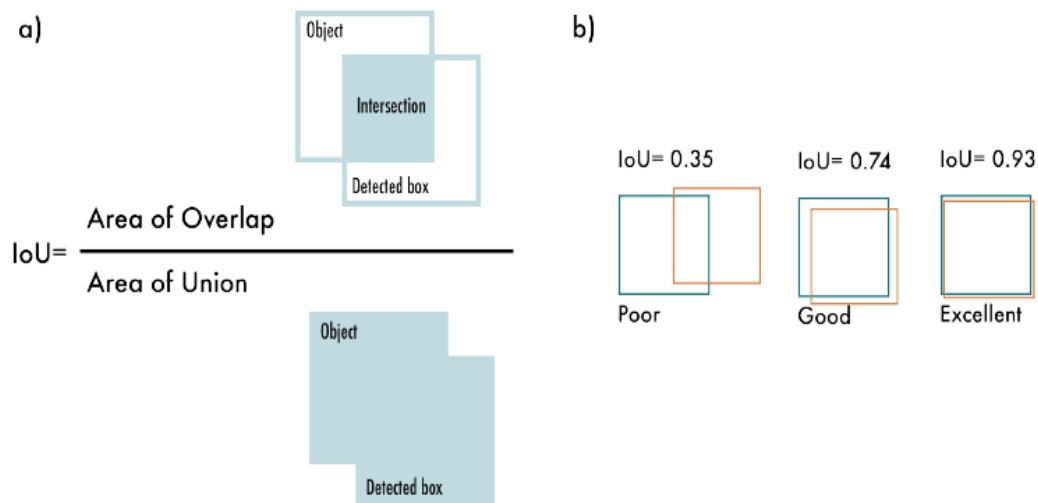


Figure 27- Intersection de la boîte englobante [54]

Comme précisé sur la figure 27, plus la superposition entre les boîtes englobantes prédites est importante, plus le score attribué à la prédiction est élevé.

C. Suppression non maximale

Lorsque plusieurs boîtes englobantes sont prédites pour un même objet, il est possible qu'elles se chevauchent ou se recouvrent partiellement. La suppression non maximale (NMS) vise à supprimer ces boîtes redondantes en conservant uniquement la boîte englobante qui a le score de confiance le plus élevé.

Le processus de NMS peut être décrit en plusieurs étapes. Tout d'abord, on sélectionne la boîte englobante ayant le score de confiance le plus élevé. Ensuite, on compare le chevauchement de cette boîte avec les autres boîtes en utilisant une mesure appelée Intersection over Union (IoU). On supprime ensuite les boîtes qui ont un chevauchement supérieur à un seuil prédéfini, généralement fixé à 50% ($IoU > 0,5$).

Après avoir effectué la suppression pour la première boîte, on passe à la boîte suivante ayant le score le plus élevé et on répète les étapes de comparaison et de suppression. Ce processus est répété pour toutes les boîtes englobantes de l'image.[55]



Figure 28 - Le résultat de Suppression non maximale [52]

➤ Version YOLO

À ce jour, YOLO a connu une évolution impressionnante avec pas moins de 8 versions déjà disponibles.



Figure 29 - Une chronologie des versions de YOLO [52]

1) YOLOv1

Depuis sa première version en 2015, YOLO a connu plusieurs améliorations et versions successives. La première version, YOLOv1, introduisait une approche novatrice en traitant la détection d'objets comme un problème de régression. Cela signifie que YOLOv1 prédit directement les boîtes englobantes et les classes d'objets dans une seule passe. Cependant, cette version avait du mal à détecter de petits objets et souffrait de problèmes de localisation précise.

2) YOLOv2

La deuxième version, YOLOv2 (également connue sous le nom de YOLO9000), a été publiée en 2016. Elle a apporté plusieurs améliorations majeures, telles que l'utilisation de l'extraction de caractéristiques résiduelles, l'ajout de différentes tailles de boîtes englobantes prédites et une meilleure gestion des objets de petite taille. De plus, YOLOv2 a introduit un mécanisme de regroupement hiérarchique, permettant de détecter un large éventail de classes d'objets.

3) YOLOv3

En 2018, YOLOv3 est sorti, apportant des améliorations supplémentaires à l'algorithme. YOLOv3 a introduit l'utilisation de plusieurs échelles de détection pour capturer des objets de différentes tailles, ainsi qu'une architecture de réseau plus profonde. Cette version a également amélioré la précision de détection des petits objets grâce à l'utilisation de filtres à différentes échelles.

4) YOLOv4

Depuis lors, d'autres variantes de YOLO ont été développées, notamment YOLOv4, était publié en 2020, a introduit des améliorations significatives en termes de vitesse et de précision, en utilisant des techniques telles que le groupement de couches, l'utilisation d'attention spatiale et la stratégie d'entraînement de l'algorithme. De plus, il a utilisé une stratégie d'augmentation de données appelée "mosaïque" lors de l'entraînement, qui combine plusieurs images en un seul échantillon d'entraînement, ce qui améliore les capacités de généralisation.

5) YOLOv5

YOLOv5, sorti la même année de YOLOv4 était publié par Glen Jocher, fondateur et PDG d'Ultralytics, s'est concentré sur l'optimisation et la simplification de l'architecture de réseau, tout en maintenant des performances de détection élevées. YOLOv5 a atteint un bon équilibre entre vitesse et précision en exploitant une architecture efficace et en mettant en œuvre des techniques avancées telles que l'entraînement et le test à plusieurs échelles, la perte focale et la détection d'objets sans ancrage.[57]

6) YOLOv6

YOLOv6 a été publié en septembre 2022 par le département Meituan Vision AI, s'est présentée des techniques de quantification améliorées utilisant la quantification après l'entraînement et la distillation au niveau des canaux, ce qui permet d'obtenir des détecteurs plus rapides et plus précis. Globalement, YOLOv6 surpasse les modèles de pointe précédents en termes de précision et de vitesse, tels que YOLOv5.[58]

7) YOLOv7

YOLOv7 a été publié sur ArXiv en juillet 2022 par les mêmes auteurs de YOLOv4 et YOLOR. À l'époque, il surpassait tous les détecteurs d'objets connus en termes de vitesse et de précision. YOLOv7 a proposé quelques modifications architecturales et une série de "bag-of-freebies" (sac de goodies), ce qui a augmenté la précision sans affecter la vitesse d'inférence, seulement le temps d'entraînement. [59]

6. Conclusion

Au cours de ce chapitre, nous avons vu que l'intelligence artificielle est un domaine en constante évolution qui offre un potentiel immense pour résoudre des problèmes complexes dans de nombreux domaines. Nous avons décrit les bases de l'apprentissage automatique et de l'apprentissage profond, ainsi que de la vision par ordinateur et des modèles de détection d'objets par CNN. Nous avons également examiné la méthode YOLO, qui est une méthode de détection d'objets rapide et précise. Ces concepts nous permettent de mieux nous préparer pour explorer les applications de l'intelligence artificielle dans des domaines sensibles tels qu'en médecine.

Chapitre 03 : L'approche proposée

1. Introduction

Dans ce chapitre, nous allons nous concentrer sur le modèle YOLOv8, une méthode de détection d'objets par CNN améliorée qui est de plus en plus utilisée dans de nombreux domaines. YOLOv8 est le dernier né de la série YOLO ; il a été introduit en Janvier 2023 par Ultralytics. Ce travail présente, à notre connaissance, la première application de YOLOv8 dans la détection et la localisation des polypes. Nous allons également aborder les datasets utilisés dans cette étude, ainsi que l'approche proposée pour améliorer les performances de la détection d'objets. Nous allons également décrire les étapes de prétraitement de données, y compris leurs annotations et augmentations, ainsi que la division de la base de données. Enfin, nous allons décrire le processus d'apprentissage du modèle YOLOv8 et les métriques utilisées pour évaluer ses performances.

2. Le modèle YOLOv8

YOLOv8 a été lancé en janvier 2023 par Ultralytics, la société qui a développé YOLOv5. C'est, à l'heure actuelle la dernière version de YOLO (You Only Look Once) qui constitue une série d'algorithmes de détection d'objets, connus pour leurs capacités de détection en temps réel [61]. YOLOv8 prend en charge plusieurs tâches de vision, telles que la détection d'objets, la segmentation, l'estimation de pose et la localisation, le suivi et la classification. [55]

Il existe cinq types dans chaque catégorie de modèles YOLOv8 : YOLOv8 Nano (YOLOv8n) est le plus rapide et ayant le minimum de paramètres, YOLOv8s est moins rapide et plus précis, puis vient YOLOv8m présente une rapidité moindre et plus de précision. YOLOv8l est encore moins rapide et encore plus précis et enfin YOLOv8 Extra Large (YOLOv8x) qui est le plus précis mais le plus lent d'entre eux.

Par rapport aux autres modèles YOLO formés à une résolution d'image de 640x640, tous les modèles YOLOv8 ont un meilleur débit avec un nombre similaire de paramètres.

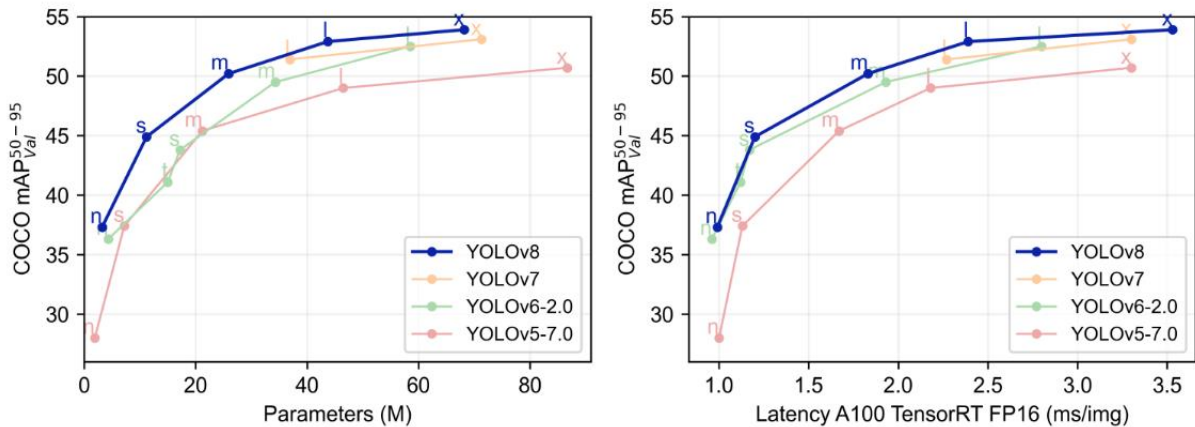


Figure 30 - YOLOv8 par rapport aux autres modèles YOLO [58]

2.1. Architecture de YOLOv8

Le modèle yolov8 utilise un réseau neuronal convolutif qui se compose de Backbone, Neck et Head, comme le montre la figure 2.

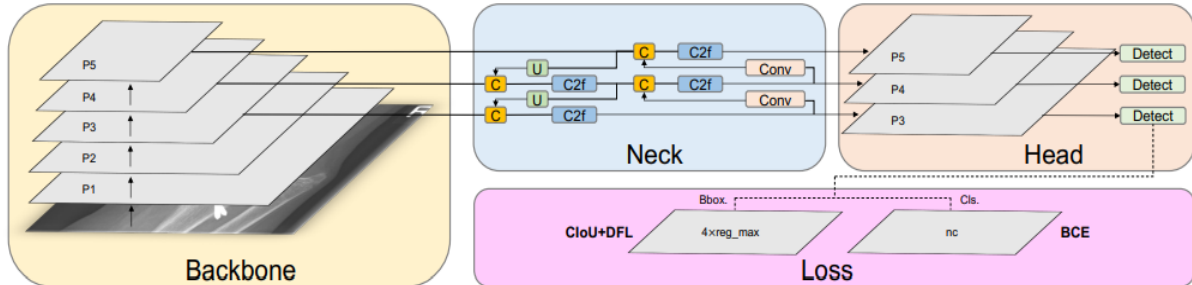


Figure 31- Architecture YOLOv8 [51]

➤ Backbone

Le modèle YOLOv8 utilise un "backbone" CSPDarknet53 modifié, qui remplace les modules Cross Stage Partial (CSP) des itérations précédentes de YOLO par des modules C2f pour l'extraction des caractéristiques. Le module C2f (2 ConvModule et n BottleNeck) combine le module C3 (3 ConvModule et n BottleNeck) et ELAN, qui a été introduit dans YOLOv7.

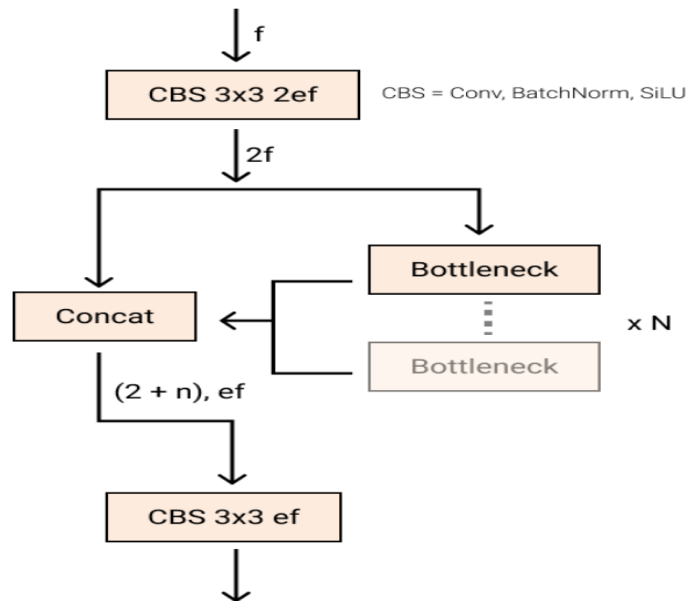


Figure 32 - Nouveau module YOLOv8 C2f

➤ Neck

La normalisation par lots et l'activation SiLU sont utilisées dans chaque couche de convolution. Le backbone se connecte au neck à trois profondeurs différentes, et le neck fusionne les caractéristiques obtenues à partir des différentes couches du réseau et les transmet à la "head".

➤ Head

Head est dissocié de sorte que les têtes de classification, de régression et de détection sont séparées. Head rend également l'algorithme YOLOv8 sans ancre en prédisant directement le centre de l'objet et les décalages des boîtes englobantes au lieu de prédire les décalages du centre par rapport aux ancres. [52]

YOLOv8 utilise une méthode de suppression non maximale douce (soft NMS) au lieu de la NMS, de sorte que les boîtes englobantes superposées ne sont pas entièrement supprimées, améliorant ainsi les performances du modèle. [51]

3. Les Datasets

Tout problème de vision par ordinateur nécessite un ensemble de données d'entraînement diversifié et de haute qualité. Les types de données disponibles pour les détecteurs de polypes colorectaux comprennent à la fois des vidéos en direct et des photos statiques extraites de flux vidéo en direct. Cette section analyse les ensembles de données standard utilisés pour construire

des modèles de détection de polypes. Dans le Tableau 1, un résumé de quelques ensembles de données.

Datasets	Résolution	Nombre d'échantillons	Résultats	Type De média	Type de tâche
Kvasir-SEG	Entre 332×487 et 1920×1072	1000	Polype	Static images	Détection, localisation & segmentation
BKAI-IGH NeoPolyp-Small	-	1200	Polype Néoplasiques & non néoplasiques	Static images	Détection & segmentation
CVC-ClinicDB	384× 288	612	Polype	Static images	Segmentation
Kvasir Capsule	-	4.820.739 images & 118 vidéos	13 classes différentes d'anomalies GI	Static images/vidéos	Classification
Coloscopic Dataset	768×576	76	Polype	Vidéo	Détection & classification

Tableau 1 - Datasets endoscopique disponible.

➤ **KVASIR-SEG**

L'ensemble de données KVASIR-SEG est une collection d'images de polypes gastro-intestinaux et de leurs masques de segmentation correspondants, annotés et vérifiés manuellement par des gastro-entérologues. L'ensemble de données se compose de 1000 images de polypes. Les images de l'ensemble de données ont des résolutions variées, allant de 332×487 à 1920×1072 pixels.

Les images et leurs masques de segmentation sont disponibles dans deux dossiers différents portant le même nom de fichier. Les images sont encodées avec une compression JPEG. Le fichier JSON contient les détails de la boîte englobante (points de coordonnées) dans les images correspondantes [55].

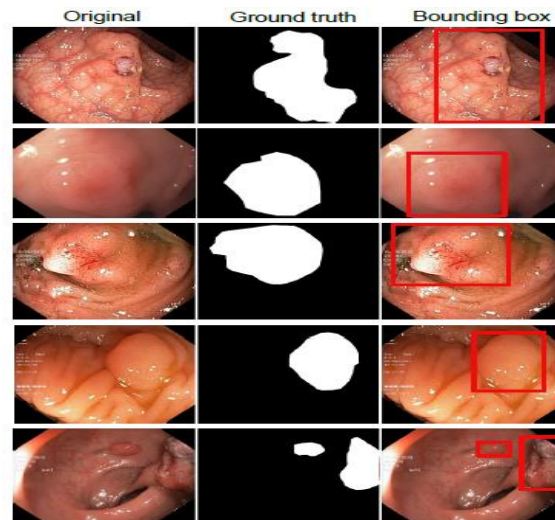


Figure 33 - Polype coloscopique, boîtes englobantes correspondantes et masques de segmentation issus de Kvasir-SEG.[52]

➤ **BKAI-IGH NeoPolyp-Small**

L'ensemble de données BKAI-IGH NeoPolyp-Small est disponible publiquement et comprend 1200 images, dont, au préalable, 1000 pour l'entraînement et 200 dans l'ensemble de test. Les polypes sont classés en néoplasiques ou adenomateux : potentiellement malin) et non néoplasiques, représentés respectivement en rouge et en vert. L'ensemble de données peut être utilisé pour la segmentation des polypes, la détection des polypes et leur localisation, ainsi que pour l'identification des caractéristiques des néoplasmes polypiques.[53]

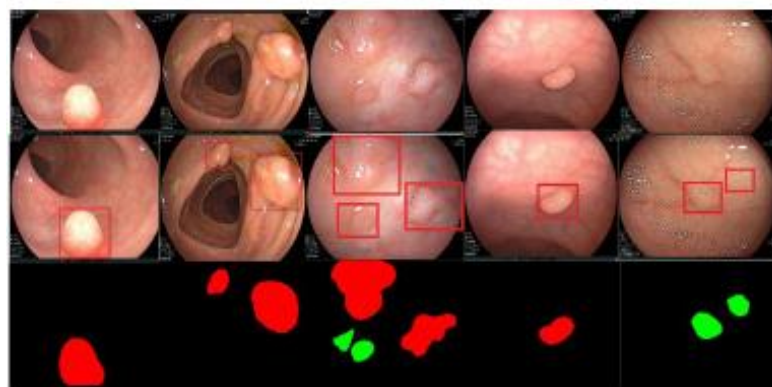


Figure 34 - Polype coloscopique, boîtes englobantes correspondantes et masques de segmentation provenant de BKAI-IGH NeoPolyp-Small [53]

4. Approche proposée

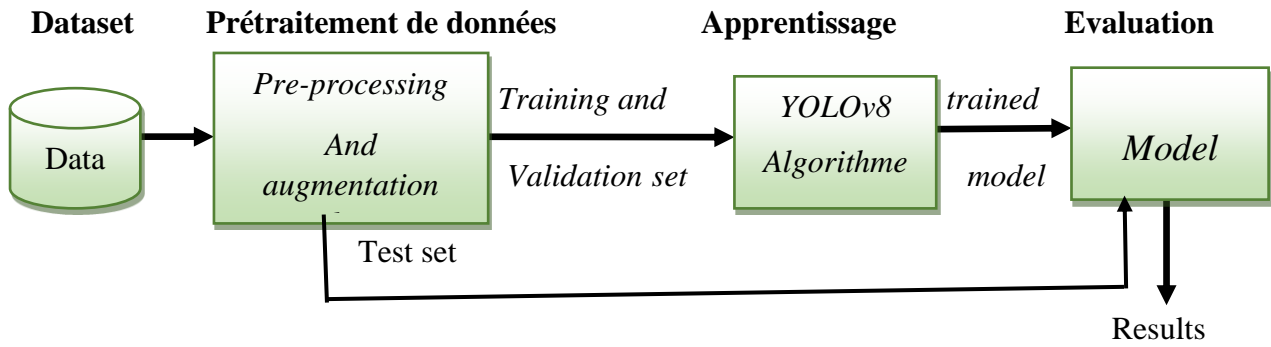


Figure 35 - Présentation de la méthodologie de la tâche de détection et de localisation des polypes

4.1. Prétraitement des données

Dans le cadre de notre étude, nous avons réalisé une combinaison des ensembles de donnée Kvasir-seg et BKAI-IGH NeoPolyp-Small. Afin de standardiser les dimensions des images pour l'entraînement du modèle, nous avons redimensionné toutes les images à une taille de 640×640 pixels.

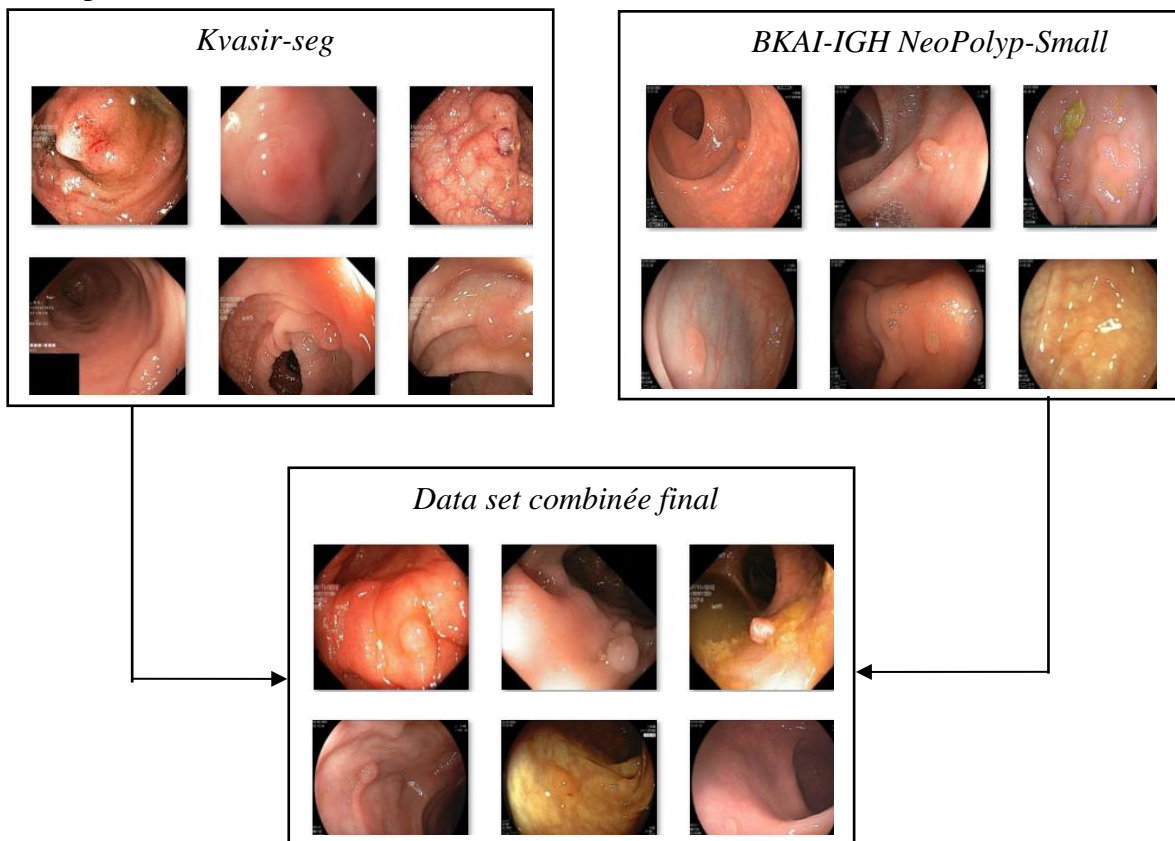


Figure 36 - données combinées pour entraîner le système de détection de polypes

4.1.1. Annotation des données

Dans notre étude, nous avons utilisé Roboflow [68] pour annoter notre ensemble de données d'images de coloscopie. À l'aide de l'interface de Roboflow, nous avons téléchargé nos images et avons procédé à l'annotation en ajoutant des boîtes englobantes précises autour de chaque polype détecté.

Cette étape d'annotation permet de fournir des informations précieuses sur la localisation et la taille des polypes dans les images. L'outil d'annotation de Roboflow offre la possibilité d'ajouter des étiquettes de la classe correspondante « polyp » pour indiquer si un polype est présent ou non dans chaque boîte englobante.

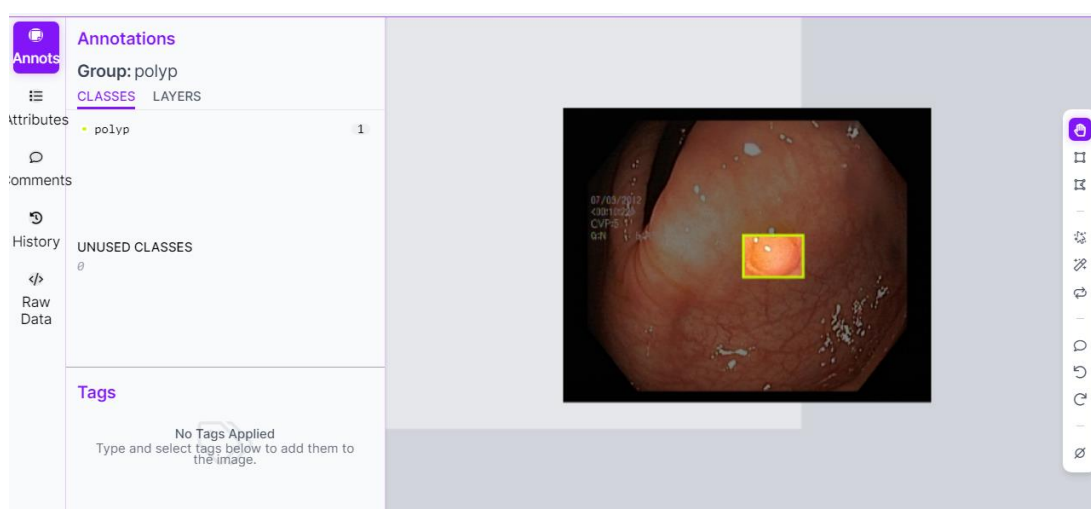


Figure 37 - exemple annotation utilisant roboflow.

Une fois les annotations terminées, nous avons enregistré les données sous le format YOLOv8, les annotations des images sont enregistrées dans un fichier (*.text), qu'il contient une ligne a le format suivant :

<Numéro de classe> <centre_x> <centre_y> <largeur> <hauteur>



Figure 38 - exemple fichier d'annotation de la base de données.

4.1.2. Augmentation des données

L'augmentation de données est une technique couramment utilisée en apprentissage automatique pour augmenter la taille du jeu de données d'entraînement en appliquant des transformations aux données existantes. Cela permet d'améliorer les performances du modèle, notamment dans les cas où la quantité de données d'entraînement disponible est limitée.

De plus, l'augmentation des données peut également aider à prévenir le sur ajustement (overfitting), où le modèle se concentre trop sur les données d'entraînement et fonctionne mal sur de nouvelles données non vues. En introduisant des variations dans les données d'entraînement, telles que la rotation ou le retournement d'images, le modèle est contraint de généraliser et d'identifier des motifs pertinents pour la tâche plutôt que de simplement mémoriser les exemples vus pendant l'entraînement.

Ci-dessous quelques techniques d'augmentation de données qui sont souvent utilisées lors de l'entraînement du modèle :

- ✓ **Flipping** : Un flip d'image consiste à faire une symétrie de l'image soit sur l'axe horizontal soit sur l'axe vertical.
- ✓ **Brightness** : La luminosité de l'image est modifiée, ce qui rendra la nouvelle image plus sombre ou plus claire
- ✓ **Zooming**: Cette augmentation effectue un zoom aléatoire sur l'image en l'agrandissant ou en la réduisant, soit en ajoutant de nouvelles valeurs de pixels autour de l'image, soit en interpolant les valeurs de pixels.
- ✓ **Noise** : Le bruit est une perturbation aléatoire ou indésirable qui peut être présente dans une image.







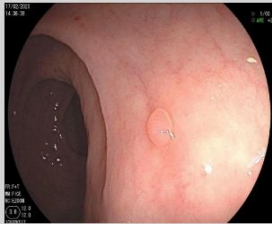

<i>Type d'augmentation</i>	<i>Image originale</i>	<i>Image augmentée</i>
<i>Vertical flip</i>		
<i>Brightness</i> <i>entre 0 et +50 %</i>		
<i>Noise</i> <i>entre 0 et +25 %</i>		
<i>Zooming</i>		

Tableau 2 - exemple d'augmentation data.

4.1.3. Division de la base de données

L'une des tâches importantes en détection d'objets consiste à identifier une méthode intelligente pour diviser l'ensemble de données en ensembles d'entraînement, de validation et de test. Dans notre étude, nous divisons la base de données en 70% pour l'apprentissage, 20% pour la validation et 10% pour le test.

4.1.4. Création du fichier de configuration du dataset

Cette étape vise à créer le fichier polyp.yaml, qui doit définir les paramètres suivants :

- **Le chemin d'entraînement (path train) :** Le chemin vers le jeu de données d'entraînement, qui devrait être un fichier YAML contenant les chemins d'image et les étiquettes associées pour chaque image.
- **Le chemin de validation (path val) :** Le chemin vers le jeu de données de validation, qui devrait également être un fichier YAML contenant les chemins d'image et les étiquettes.
- **Le chemin de test (path test) :** Le chemin vers le jeu de données de test, qui devrait également être un fichier YAML contenant les chemins d'image et les étiquettes.
- **Le nombre de classes (cn) :** Dans notre cas, nc=1 car la seule classe d'objet est "polyp".
- **La liste des noms (ln) :** Une liste des noms de classe, qui pour notre dataset serait ['polyp'].

```
train: /content/drive/MyDrive/asma-2023/dataset-asma/train/images
val: /content/drive/MyDrive/asma-2023/dataset-asma/valid/images
test: /content/drive/MyDrive/asma-2023/dataset-asma/test/images

nc: 1
names: ['polyp']
```

Figure 39 - Contenu du fichier polyp.yaml pour YOLOv8.

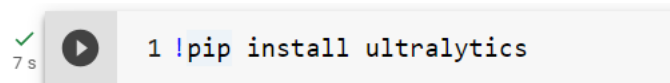
4.2. Apprentissage de modèle YOLOv8

Dans cette étude, nous avons entraîné un modèle YOLOv8 nano (YOLOv8n) à partir de zéro. Le modèle YOLOv8n a obtenu une précision moyenne de validation (mAP) de 37,3 sur l'ensemble de données COCO val2017. Il a été exécuté à une vitesse de 0,99 milliseconde par image sur un GPU NVIDIA A100 en utilisant des poids exportés en TensorRT. Nous avons choisi ce modèle pour sa rapidité de prédiction afin de la détection et localisation des polypes [66].

Les étapes d'entraînement du modèle sont comme suivies :

1. Téléchargement du modèle YOLOv8n :

Cette étape consiste à connecter google Colab avec notre google drive, ensuite installer la bibliothèque Ultralytics à partir de Pip qu'il contient le modèle yolov8 par la commande suivante :



```
1 !pip install ultralytics
```

Ensuite, nous configurons notre environnement d'entraînement fourni par Google Colab pour qu'il fonctionne avec le GPU.

Le GPU nous permettra d'accélérer le temps d'entraînement. Colab est également pratique car il est préinstallé avec Torch et CUDA.

Après la configuration de colab, en utilisant le lien vers nos données annotées, et nous téléchargeons directement les données dans Colab au format YOLOv8 PyTorch.

2. Préparer le model pour l'apprentissage :

Dans cette étape nous définissons la configuration du modèle YOLOv8 :

D'abord, nous choisissons le modèle de base le plus petit et le plus rapide de YOLOv8, qui est YOLOv8n. Ce modèle a été choisi pour sa vitesse de prédiction en temps réel [63].

Ensuite, nous définissons les paramètres spécifiques du modèle YOLOv8 en fonction de nos besoins. Voici les paramètres que nous avons choisis :

- La taille de l'image : Nous spécifions la taille de l'image à 640x640 pixels. Cela signifie que les images utilisées pour l'entraînement seront redimensionnées à cette taille avant d'être fournies au modèle.
- Le batch size : Nous définissons le batch size à 16. Cela signifie que 16 images seront traitées simultanément lors de chaque itération d'entraînement.
- Le nombre d'époques : Nous définissons le nombre d'époques à 27. Cela indique le nombre de fois où l'ensemble d'entraînement complet sera parcouru par le modèle lors de l'entraînement.

De plus, nous devons fournir le chemin vers le fichier "dataset.yaml" dans le paramètre "data", Ce fichier contient les informations sur le jeu de données utilisé pour l'entraînement, y compris les chemins vers les images, les annotations, les classes, etc...

4.3. Métriques d'évaluation des performances

L'évaluation des performances d'un modèle d'objet se fait à l'aide d'indicateurs calculés à partir de la matrice de confiance. Cette matrice est composée de chiffres qui permettent d'obtenir une vision globale de l'efficacité du modèle sont :

- ✓ **Vrai positif (TP)** : le nombre de fois où le modèle prédit correctement la classe positive.
- ✓ **Faux positif (FP)** : le nombre de fois où le modèle prédit la classe positive alors que le résultat attendu était la classe négative.
- ✓ **Vrai négatif (TN)** : le nombre de fois où le modèle prédit correctement la classe négative.
- ✓ **Faux négatif (FN)** : le nombre de fois où le modèle prédit la classe négative alors que le résultat attendu était la classe positive.

Exemple :

<i>Vrai positif</i> Prédiction : Polype Résultat attendu : Polype	<i>Faux positif</i> Prédiction : Polype Résultat attendu : pas de Polype
<i>Faux négatif</i> Prédiction : pas de Polype Résultat attendu : Polype	<i>Vrai négatif</i> Prédiction : pas de Polype Résultat attendu : pas de Polype

Tableau 3 - comparaison entre les vrai et les faux positive et négative.

Ces indicateurs sont utilisés pour mesurer la précision, le rappel, la précision moyenne (mAP) et le F1-score du modèle [45].

1) La précision

La précision est une mesure de l'exactitude des prédictions pour identifier uniquement les objets pertinents. Elle est calculée en tant que rapport de tous les Vrais Positifs sur l'ensemble des détections effectuées par le modèle.

$$precision = \frac{TP}{TP+FP}$$

2) le rappel

C'est pour calculer les prédictions correctes parmi toutes les données correctement prédites, on utilise le terme de "vrais positifs". Il est calculé en tant que rapport de tous les Vrais Positifs sur l'ensemble des vérités terrain.

$$\text{rappel} = \frac{TP}{TP+FN}$$

3) *F1-score*

Le score F1 est calculé comme la moyenne harmonique de la précision et du recall, en tenant compte à la fois des faux positifs et des faux négatifs. La formule pour calculer le score F1 est la suivante :

$$F1\text{-score} = 2 \times \frac{(\text{precision} \times \text{recall})}{\text{precision} + \text{recall}}$$

4) L'accuracy

Accuracy est une métrique de performance qui représente la proportion ou la fraction de prédictions correctes effectuées par le modèle par rapport au nombre total de prédictions. La formule pour calculer la précision est la suivante :

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

5) Le mAP (Mean Average Precision)

Le mAP (Mean Average Precision) est une mesure de la capacité d'un système à récupérer les éléments pertinents en réponse à une requête de l'utilisateur. Il est calculé en prenant la moyenne des scores de précision à différents niveaux de recall :

$$mAP = \sum_{k=1}^{k=n} AP_k$$

5. Conclusion

Ce chapitre nous a permis de nous concentrer sur le modèle YOLOv8, une méthode de détection d'objets par CNN améliorée qui est de plus en plus utilisée dans de nombreux domaines. Nous avons également décrit les datasets utilisés dans cette étude, ainsi que l'approche proposée pour améliorer les performances de la détection d'objets. Nous avons également examiné les étapes de prétraitement de données, y compris leurs annotations et leurs augmentations. Nous avons abordé la notion de division de la base de données entre apprentissage, test et validation. Enfin, nous avons décrit le processus d'apprentissage du modèle YOLOv8 et les métriques utilisées pour évaluer ses performances. Ces éléments clés nous permettent d'aborder la phase d'implémentation et d'expérimentation dans la détection d'objets par CNN.

Chapitre 04 : L'implémentation et les résultats obtenus

1. Introduction

La phase d'implémentation est une étape cruciale au terme de laquelle le système informatique est mis en œuvre à l'aide d'outils informatiques adéquats et appropriés. Le produit logiciel, ainsi obtenu, doit être convivial, ergonomique et facile à utiliser par des non-informaticiens et être conforme aux cahiers des charges et spécifications et attentes des utilisateurs finaux.

Nous décrivons, dans ce chapitre, les étapes de mise en œuvre de notre application informatique.

2. Le matériel utilisé

Notre configuration matérielle inclut les dispositifs suivants :

- Modèle : HP ProBook 450 G3.
- Processeur : Intel(R) Core (TM) i5-6200U CPU@ 2.30GHz.
- Mémoire : 8.00 Go.
- Système d'exploitation : Windows 10, 64 bits.

3. Environnements et outils de développement utilisés

➤ Google colab



Figure 40- logo de Google colaboratory

Google Colab [69] est une plateforme de notebooks Jupyter hébergée sur le cloud par Google. Elle permet aux utilisateurs d'écrire, d'exécuter et de partager du code Python de manière interactive.

Avec Google Colab, il est possible de créer et d'exécuter des fichiers texte (notebook) contenant du code Python, des visualisations, des textes explicatifs, des graphiques et autres. L'environnement Colab est entièrement géré par Google, ce qui signifie que l'on n'a pas besoin d'installer ou de configurer de logiciels supplémentaires. De plus, Colab offre l'avantage d'accéder gratuitement à des ressources de calcul puissantes de Google, y compris les processeurs graphiques (GPU) et les unités de traitement tensoriel (TPU). Cela en fait un outil populaire pour l'apprentissage automatique, l'analyse de données et la création de prototypes de projets.



Figure 41 - L'environnement Google Colaboratory

Colab est très largement utilisé par la communauté machine learning comme c'est le cas dans les applications suivantes :

- TensorFlow.
- Développement et entraînement de réseaux de neurones.
- Expérimentation avec les TPU.
- Dissémination de la recherche en IA.
- Création de tutoriels.

➤ **Python**



Figure 42 - Logo de python.

Python est un langage de programmation de haut niveau, polyvalent et très populaire. Le langage de programmation Python (le dernier Python 3) est utilisé dans le développement Web, les applications d'apprentissage automatique, ainsi que dans toutes les technologies de pointe de l'industrie logicielle. Le langage de programmation Python convient très bien aux débutants, ainsi qu'aux programmeurs expérimentés dans d'autres langages de programmation comme C++ et Java [70].

➤ **Roboflow.**



Figure 43 - Logo de roboflow.

Roboflow est une plateforme de vision par ordinateur lancée en janvier 2020. Elle permet aux utilisateurs de créer des modèles de vision par ordinateur plus rapidement et plus précisément grâce à de nouvelles techniques de collecte de données, de prétraitement et d'entraînement des modèles. Roboflow permet aux utilisateurs de télécharger des ensembles de données personnalisés, de dessiner des annotations, de modifier les orientations des images, de redimensionner les images, de modifier le contraste des images et d'effectuer des augmentations de données. Il peut également être utilisé pour entraîner des modèles.

Roboflow prend en charge les formats d'annotation les plus populaires, notamment JSON, XML, CSV et TXT. Roboflow dispose également d'un outil de conversion d'annotations universel qui permet aux utilisateurs de télécharger et de convertir des annotations d'un format à un autre sans avoir à écrire de scripts de conversion pour des ensembles de données de détection d'objets personnalisés. [71]

➤ **Opencv**



Figure 44 - Logo OpenCV

OpenCV (Open Source Computer Vision Library) est une bibliothèque logicielle open source de vision par ordinateur et d'apprentissage automatique. OpenCV a été construit pour fournir une infrastructure commune pour les applications de vision par ordinateur et pour accélérer l'utilisation de la perception artificielle dans les produits commerciaux. Étant un produit sous licence BSD, OpenCV permet aux entreprises d'utiliser et de modifier facilement le code.[72] La bibliothèque compte plus de 2500 algorithmes optimisés, ce qui inclut un ensemble complet d'algorithmes de vision par ordinateur et d'apprentissage automatique classiques et de pointe. Ces algorithmes peuvent être utilisés pour détecter et reconnaître des visages, identifier des objets, classer des actions humaines dans des vidéos, suivre les mouvements de la caméra, suivre des objets en mouvement, extraire des modèles 3D d'objets, etc. [72]

➤ **Tkinter**



Figure 45 - Logo tkinter.

Tkinter est une bibliothèque d'interface utilisateur graphique (GUI) open source et portable conçue pour être utilisée dans les scripts Python. Tkinter s'appuie sur la bibliothèque Tk, la bibliothèque graphique utilisée par Tcl/Tk et Perl, qui est à son tour implémentée en C. Par conséquent, on peut dire que Tkinter est implémenté en utilisant plusieurs couches. Plusieurs kits d'outils GUI concurrents sont disponibles pour être utilisés avec le langage Python. [73]

➤ PyCharm



Figure 46 - Logo PyCharm.

PyCharm est un environnement de développement intégré utilisé pour programmer en Python. Il permet l'analyse de code et contient un débogueur graphique. Il permet également la gestion des tests unitaires, l'intégration de logiciel de gestion de versions, et supporte le développement web avec Django.

Développé par l'entreprise tchèque JetBrains, c'est un logiciel multi-plateforme qui fonctionne sous Windows, Mac OS X et GNU/Linux. Il est décliné en édition professionnelle, diffusé sous licence propriétaire, et en édition communautaire diffusé sous licence Apache.[74]

4. L'implémentation

Afin d'offrir une plateforme d'exploitation ergonomique et assez conviviale englobant un maximum de fonctionnalités concernant le travail entrepris, nous avons développé une application logicielle qui permet à l'utilisateur de consulter le système de façon interactive. Les interfaces principales de cette application sont présentées de façon brève ci-dessous.

4.1. Les interfaces graphiques de l'application

Dans cette partie nous allons présenter les différentes parties de notre système de détection des polypes. Nous allons présenter les interfaces principales qui offrent un aperçu des fonctionnalités de l'application.



Figure 47 - L'interface principale de l'application.

Notre application utilise le modèle de détection et de localisation des polypes dans une image ou un enregistrement vidéo dont nous avons fait l'apprentissage.



Figure 48 - Les tâches de l'application.

➤ **La détection et la localisation des polypes dans une image**

Pour que le système puisse détecter des polypes dans une image, on appuie sur le bouton « importe une image » pour choisir une image (Plusieurs formats d'image sont acceptés tels que les extensions des images sont .jpg, png). La figure 50 ci-dessous montre cette interface.

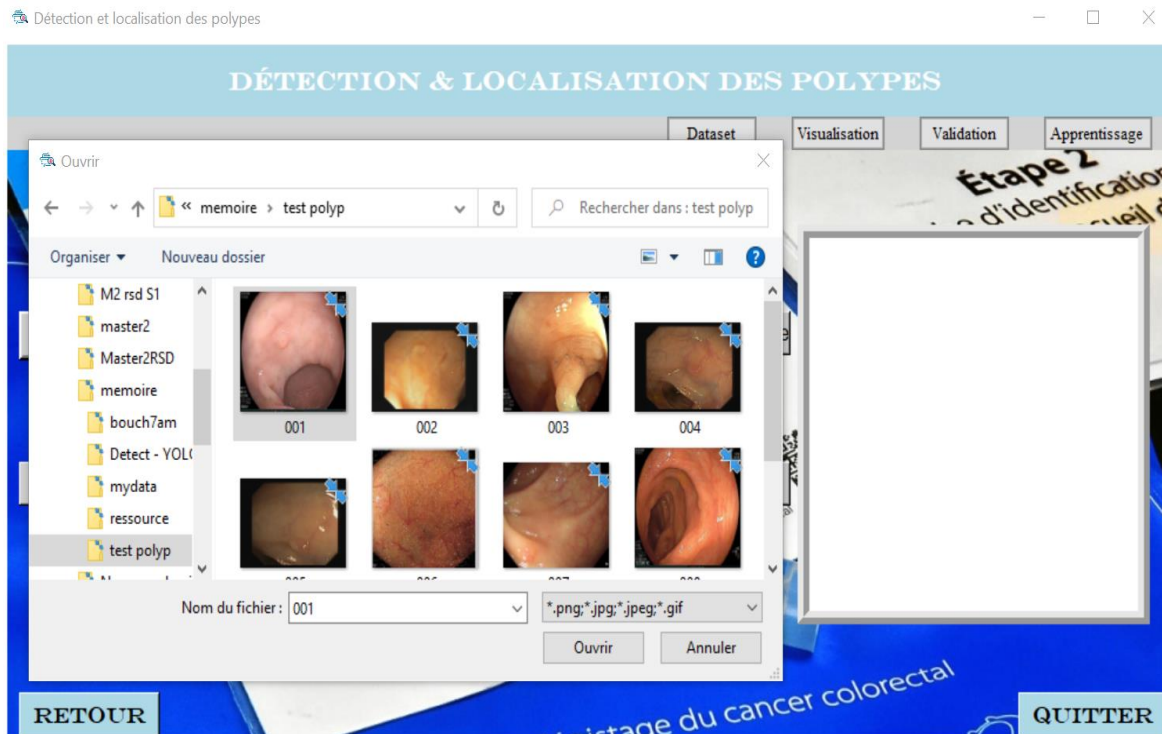


Figure 49 - Le choix d'une image.

➤ **La détection et la localisation des polypes dans une vidéo**

Pour que le système puisse détecter des objets dans une vidéo, il faut choisir le bouton « importe une vidéo ». La figure 51 ci-dessous illustre ce principe. Puis, le système va extraire les images et changer leurs tailles selon le format utilisé (300×300 pour l'application sous test).

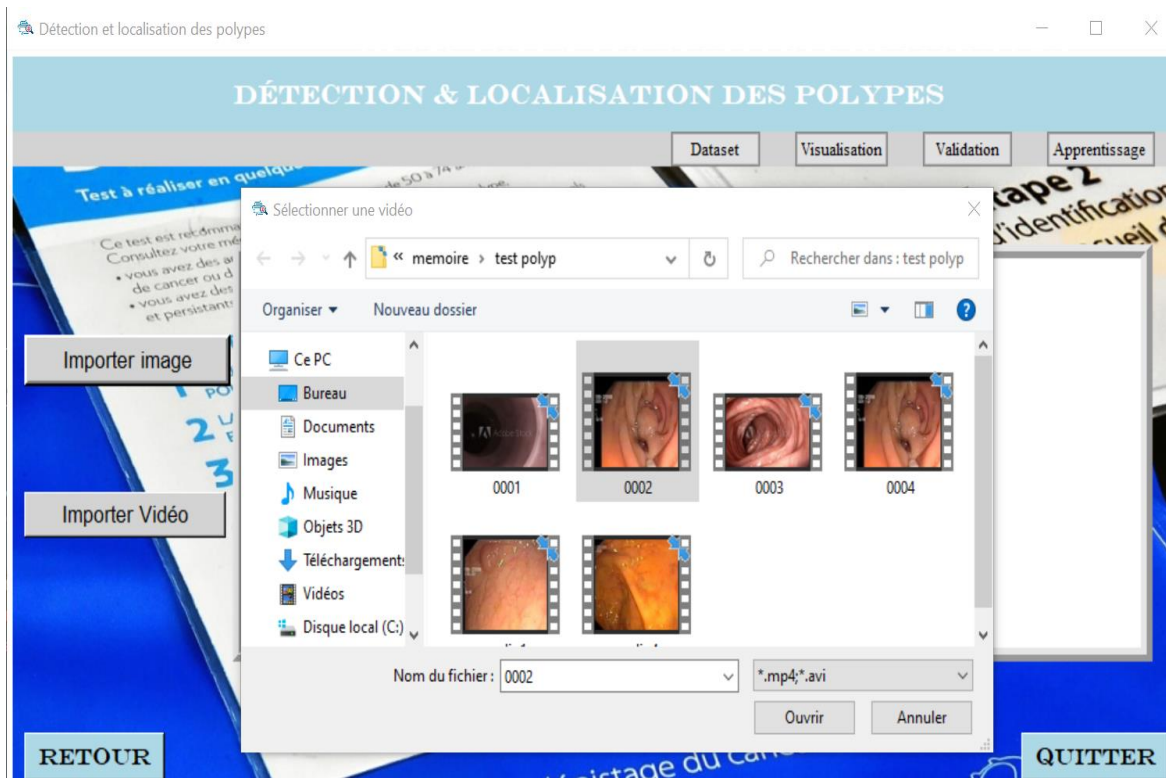


Figure 50 - Le choix d'un enregistrement vidéo.

5. Les résultats obtenus

5.1. Apprentissage

Le modèle YOLOv8n est composé de 225 couches, 3157200 paramètres et 3157184 gradients. Le modèle a été entraîné en utilisant une puissance de calcul de 8.2 GFLOPs. La figure 52 illustre le modèle d'entraînement.

```

Overriding model.yaml nc=80 with nc=1

      from n  params module                        arguments
  0         -1 1    464  ultralytics.nn.modules.Conv                    [3, 16, 3, 2]
  1         -1 1   4672 ultralytics.nn.modules.Conv                    [16, 32, 3, 2]
  2         -1 1   7360 ultralytics.nn.modules.C2f                     [32, 32, 1, True]
  3         -1 1  18560 ultralytics.nn.modules.Conv                    [32, 64, 3, 2]
  4         -1 2  49664 ultralytics.nn.modules.C2f                     [64, 64, 2, True]
  5         -1 1  73984 ultralytics.nn.modules.Conv                    [64, 128, 3, 2]
  6         -1 2 197632 ultralytics.nn.modules.C2f                     [128, 128, 2, True]
  7         -1 1 295424 ultralytics.nn.modules.Conv                    [128, 256, 3, 2]
  8         -1 1 460288 ultralytics.nn.modules.C2f                     [256, 256, 1, True]
  9         -1 1 164608 ultralytics.nn.modules.SPPF                    [256, 256, 5]
 10        -1 1     0  torch.nn.modules.upsampling.Upsample        [None, 2, 'nearest']
 11        [-1, 6] 1     0  ultralytics.nn.modules.Concat                 [1]
 12        -1 1 148224 ultralytics.nn.modules.C2f                     [384, 128, 1]
 13        -1 1     0  torch.nn.modules.upsampling.Upsample        [None, 2, 'nearest']
 14        [-1, 4] 1     0  ultralytics.nn.modules.Concat                 [1]
 15        -1 1   37248 ultralytics.nn.modules.C2f                     [192, 64, 1]
 16        -1 1   36992 ultralytics.nn.modules.Conv                    [64, 64, 3, 2]
 17        [-1, 12] 1     0  ultralytics.nn.modules.Concat                 [1]
 18        -1 1 123648 ultralytics.nn.modules.C2f                     [192, 128, 1]
 19        -1 1 147712 ultralytics.nn.modules.Conv                    [128, 128, 3, 2]
 20        [-1, 9] 1     0  ultralytics.nn.modules.Concat                 [1]
 21        -1 1 493056 ultralytics.nn.modules.C2f                     [384, 256, 1]
 22        [15, 18, 21] 1 751507 ultralytics.nn.modules.Detect                 [1, [64, 128, 256]]

Model summary: 225 layers, 3011043 parameters, 3011027 gradients, 8.2 GFLOPs
    
```

Figure 51 - Résumé du modèle YOLOv8 lors de l'entraînement du modèle.

Nous avons utilisé l'optimiseur Stochastic Gradient Descent (SGD) pour l'entraînement du modèle yolov8n.pt. Le temps nécessaire pour entraîner le modèle YOLOv8n sur 8232 images dans l'ensemble de données d'entraînement et 934 images dans l'ensemble de données de validation était de 3,238 heures.

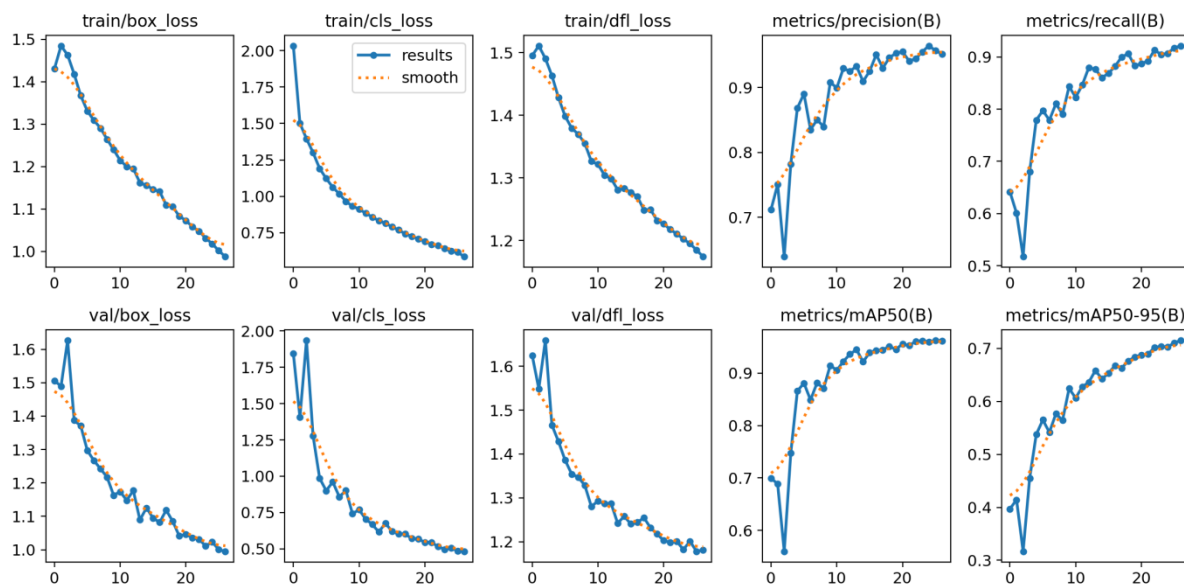


Figure 52 - Graphique des performances de l'entraînement du modèle YOLOv8.

La figure 53 affiche les graphiques pour la perte de boîte englobante (box_loss), la perte de classification (cls_loss) et la perte de la couche de convolution déformable (dfl_loss) pour les ensembles d'entraînement et de validation.

- **La perte dans la boîte englobante lors de la prédiction :** est calculée comme étant l'erreur entre la boîte englobante prédite et la boîte englobante réelle. Cela signifie en fait que plus la valeur de la perte de boîte est faible, plus l'algorithme est précis dans la prédiction de la boîte englobante ayant un chevauchement maximal avec la boîte englobante réelle.
- **La perte de la couche de convolution déformable :** est une nouvelle addition à l'architecture YOLO dans YOLOv8. Cette perte mesure l'erreur dans les couches de convolution déformables, qui sont conçues pour améliorer la capacité du modèle à détecter des objets avec différentes échelles et ratios d'aspect. Une valeur plus faible de "dfl_loss" indique que le modèle est meilleur pour gérer les déformations d'objets et les variations d'apparence.

- **La perte de classification** : est mesure l'erreur dans les probabilités de classe prédites pour chaque objet dans l'image par rapport à la vérité terrain. Une valeur plus faible de "cls_loss" signifie que le modèle prédit plus précisément la classe des objets.

Le modèle a rapidement montré des valeurs améliorées pour la précision, le rappel (recall) et le mAP après environ 20 époques. De plus, la perte de boîte englobante et la perte d'objet ont montré une baisse rapide après environ 10 époques.

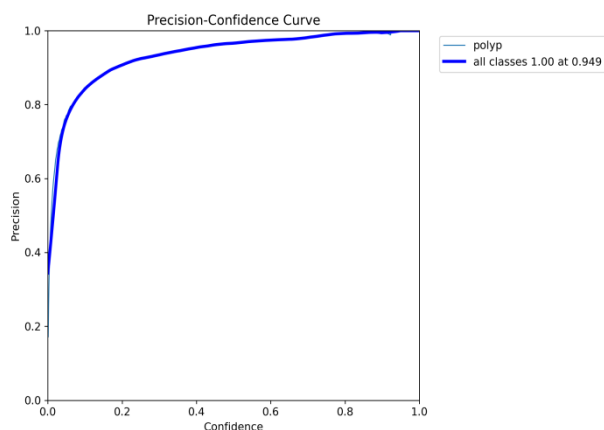


Figure 53 - la courbe de précision.

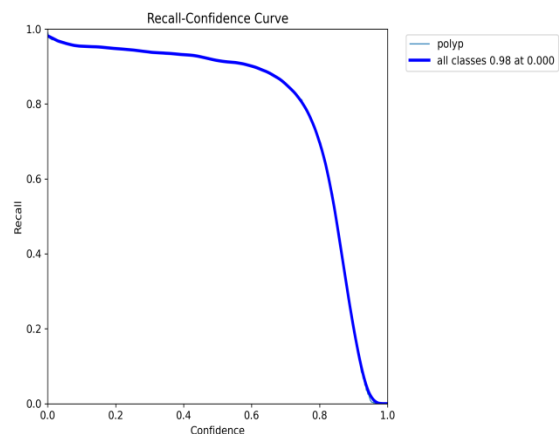


Figure 54 - La courbe du rappel.

Dans la Figure 54, on observe que lorsque le seuil de confiance augmente, la précision du détecteur de polype augmente également. Cela indique que le détecteur devient plus précis dans ses prédictions, mais il peut également manquer certains objets positifs, ce qui entraîne une diminution du rappel. En revanche, dans la Figure 55, on constate que lorsque le seuil augmente, le rappel diminue progressivement, ce qui signifie que le détecteur manque de plus en plus d'objets positifs.

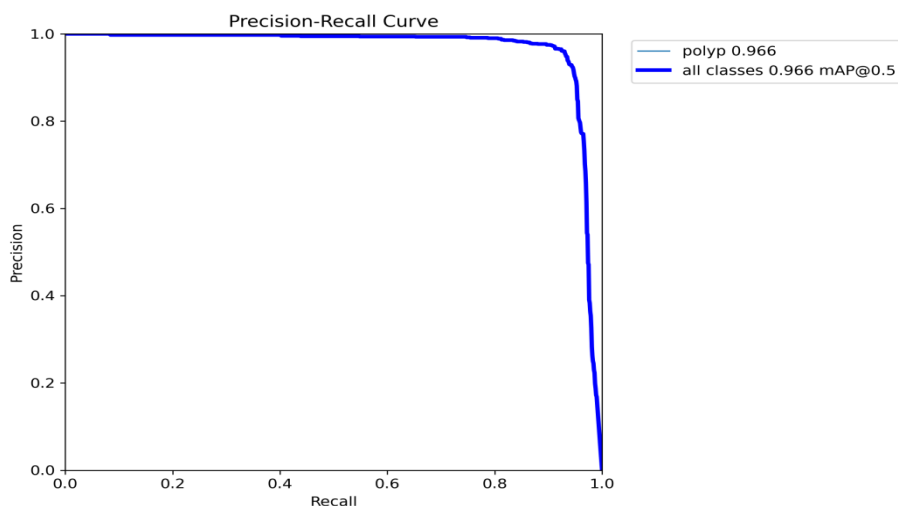


Figure 55 - La courbe PR (Precision-Recall) pour l'entraînement du modèle.

La Figure 56 montre la courbe PR pour le modèle YOLOv8 entraîné sur l'ensemble de données. Le graphique montre un bon compromis entre les valeurs de précision et de rappel, indiquant une bonne surface sous la courbe PR avec une valeur mAP@0.5 de 0,966.

5.2. Evaluation du modèle

Pour s'assurer que le modèle YOLOv8 généralise bien aux nouvelles données, les performances du modèle ont été évaluées sur un ensemble de validation après chaque époque d'entraînement.

Le tableau 4 illustre les résultats de performance obtenus à partir de l'évaluation du modèle :

Modèle	mAP_50	Precision %	Recall %	F1_score
YOLOv8	96.63%	96.00%	92.92%	94.43%

Tableau 4- Résultats des performances du modèle YOLOv8.

D'excellents résultats ont été obtenus sur les mesures de précision de détection des polypes en utilisant YOLOv8. Une précision moyenne de 96,63% (mAP_50) sur l'ensemble de données a été atteinte. Plus précisément, on a obtenu une précision de 96% et un rappel de 92,92%, ce qui correspond à un score F1 de 94,43%. La figure 57 présente la matrice de confusion utilisée.

Ces résultats montrent que le modèle YOLOv8 est capable de détecter les objets avec une grande précision et un rappel élevé, ce qui en fait un choix solide pour les tâches de détection et localisation des polypes.

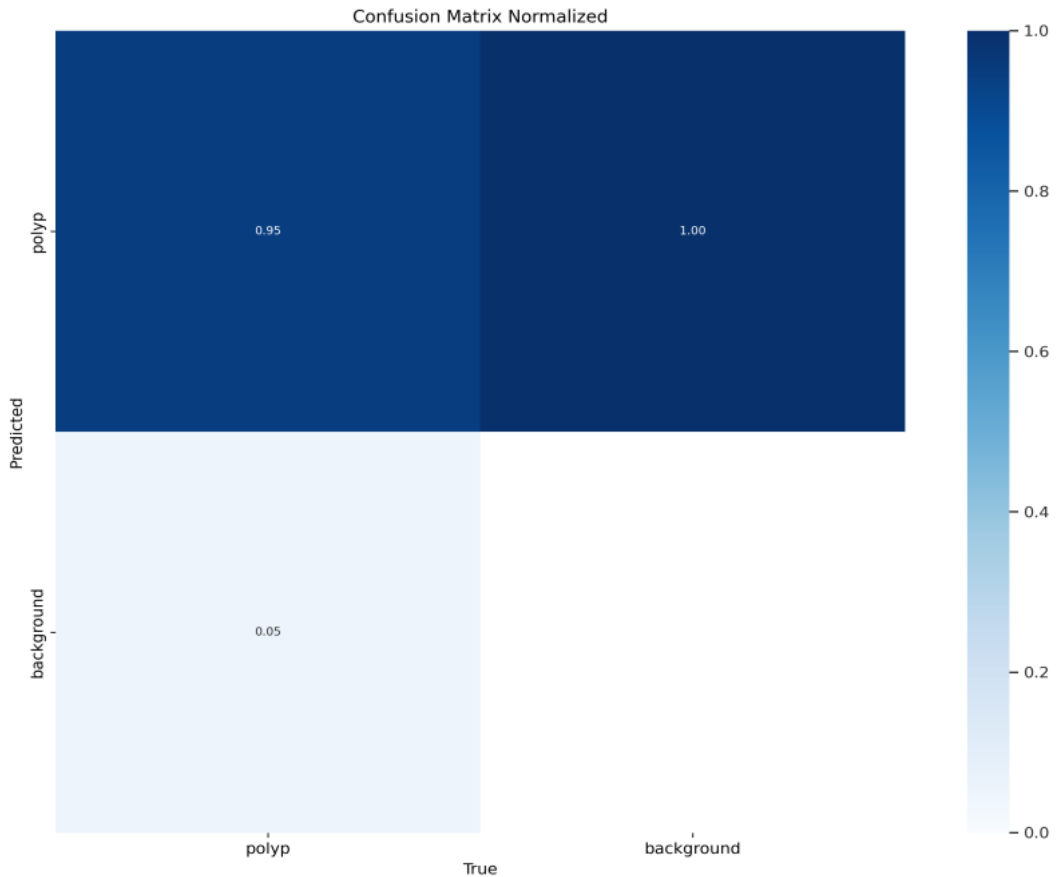


Figure 56- la matrice de confusion de la classe polype.

La figure 58 montre les étiquettes de lot de validation et leurs prédictions correspondantes dans des boîtes de couleur rouge réalisées lors de l'évaluation du modèle YOLOv8 avec l'ensemble de données de base d'entraînement. Les étiquettes réelles indiquent les classes réelles des objets présents dans l'image, tandis que les prédictions sont les classes prédites par le modèle pour ces objets.

Cela permet de comparer visuellement les prédictions du modèle avec les étiquettes réelles, afin d'évaluer ses performances. De plus aidé à comprendre les erreurs de prédiction spécifiques, telles que les faux positifs (objets incorrectement identifiés) ou les faux négatifs (objets manqués par le modèle).

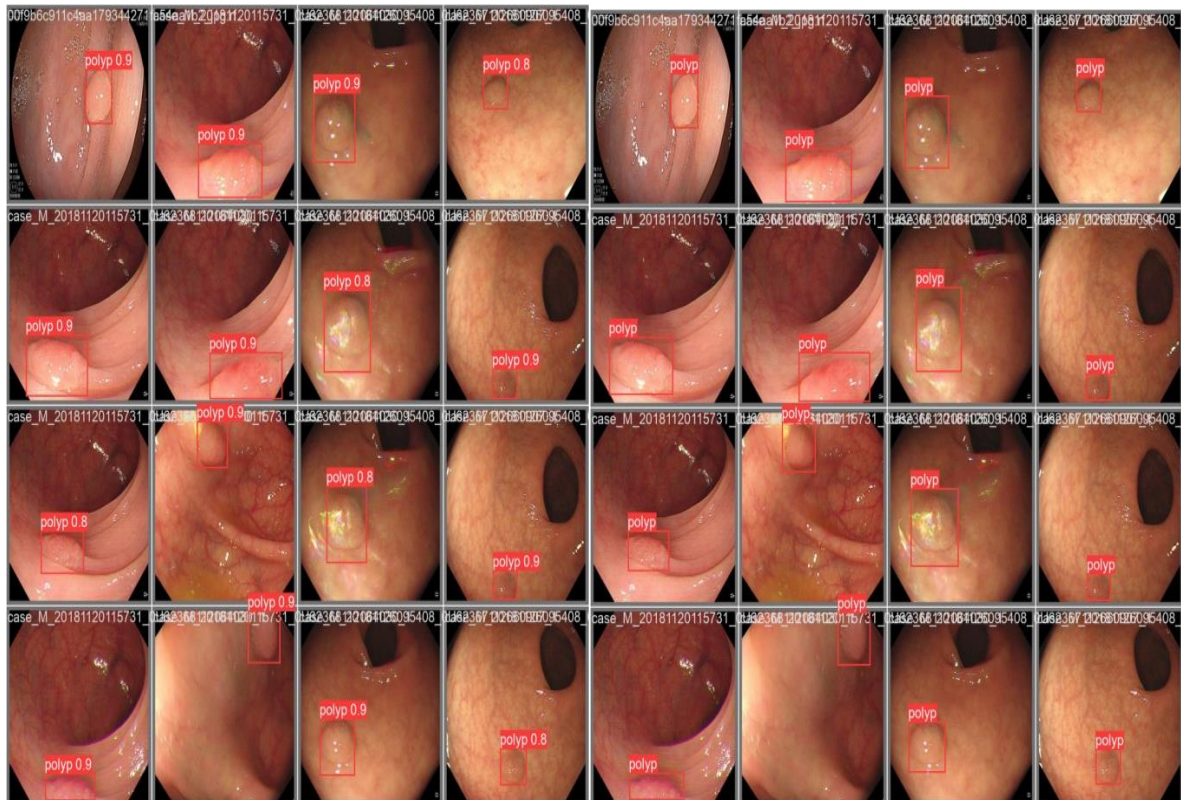
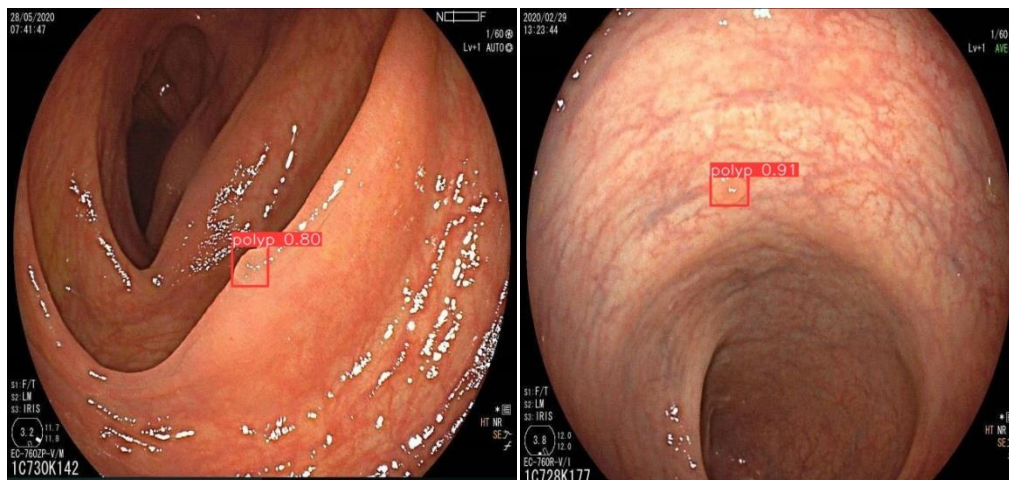


Figure 57 - Étiquettes et prédictions correspondantes pour la validation par lots (batch).

5.3. Test et résultats

Après avoir validé le modèle et pour évaluer leur performance nous avons fait des tests sur l'ensemble de données du test, et on a obtenu les résultats suivants :

➤ **La détection des polypes dans les images**



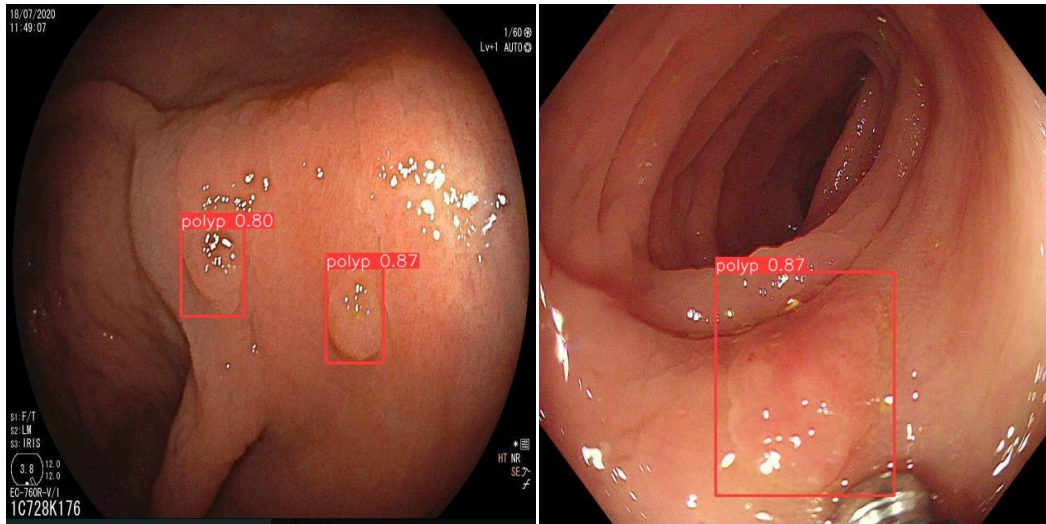


Figure 58 - Quelques résultats de tests de détection obtenus par YOLOv8.

Les résultats obtenus (Figure 59) lors de la phase de test ont été extrêmement encourageants. Le modèle a réussi à détecter et localiser avec précision la présence des polypes dans les images. Ces tests approfondis confirment que le modèle de détection de polypes est robuste et fiable, ouvrant ainsi la voie à des applications cliniques potentielles pour la détection précoce et le traitement des maladies gastro-intestinales liées aux polypes.

➤ **La détection des polypes dans les vidéos :**

Le modèle a également été testé sur un ensemble de données vidéo afin de visualiser le comportement du modèle lorsque l'entrée est une vidéo d'une coloscopie pour le modèle YOLOv8.

```
1 %cd /content/drive/MyDrive/Asma
2 model = YOLOv8 ("content/drive/MyDrive/Asma/runs/detect/train/weights/best.pt")
3 results = model.predict(source="content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4", conf=0.25 ,save=True)

video 1/1 (497/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 7.9ms
video 1/1 (498/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.1ms
video 1/1 (499/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 8.5ms
video 1/1 (500/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.8ms
video 1/1 (501/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.5ms
video 1/1 (502/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 8.3ms
video 1/1 (503/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.4ms
video 1/1 (504/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.4ms
video 1/1 (505/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 9.0ms
video 1/1 (506/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 8.3ms
video 1/1 (507/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.3ms
video 1/1 (508/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.5ms
video 1/1 (509/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.4ms
video 1/1 (510/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.5ms
video 1/1 (511/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.3ms
video 1/1 (512/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 17.2ms
video 1/1 (513/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.1ms
video 1/1 (514/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 7.2ms
video 1/1 (515/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.6ms
video 1/1 (516/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.9ms
video 1/1 (517/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 11.2ms
video 1/1 (518/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.6ms
video 1/1 (519/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 5.4ms
video 1/1 (520/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 9.2ms
video 1/1 (521/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 6.3ms
video 1/1 (522/552) /content/drive/MyDrive/Asma/Polypectomie_anse_2.mp4: 544x640 1 polyp, 8.0ms

Speed: 2.6ms preprocess, 6.9ms inference, 1.1ms postprocess per image at shape (1, 3, 640, 640)
Results saved to runs/detect/predict5
```

Figure 59 - Détection des polypes dans une vidéo par YOLOv8.

La figure 60 présente l'exécution du code pour la détection sur un fichier vidéo de l'examen coloscopique à l'aide du modèle YOLOv8 entraîné. Lors de l'exécution, la vidéo est traitée image par image. Le modèle considère chaque image comme une donnée d'entrée pour effectuer la tâche de détection des polypes. Les résultats de la détection sont enregistrés sous forme d'une nouvelle vidéo où les détails des boîtes encadrant les polypes détectés sont mis en évidence pour chaque image de la séquence vidéo.

L'exécution du traitement vidéo a été réalisée à une vitesse de 94,33 images par seconde (FPS), Cela indique une vitesse de traitement très élevée, ce qui est bénéfique pour les applications nécessitant un traitement rapide en temps réel.

6. Etudes comparatives

Nous avons effectué une évaluation comparative de YOLOv8 avec un autre modèle de détection d'objets, YOLOv7, sur un ensemble de données que nous avons créé. Les résultats présentés

dans le tableau 1 montrent que le modèle YOLOv8 a surpassé l'autre modèle en termes de détection et de reconnaissance d'objets, avec moins d'erreurs.

Comme les autres modèles, YOLOv8 a également été formé sur une fonction de perte. Cependant, il se distingue comme le modèle de détection d'objets à usage général le plus rapide disponible, et il représente actuellement l'état de l'art en matière de détection d'objets

Models	mAP_50 %	Precision %	Recall %	F1_score %
YOLOv7	83.72%	86.54%	79.98%	83.13%
YOLOv8	96.63%	96%	92.92%	94.43%

Tableau 5- Résultats de comparaison entre YOLOv8 et YOLOv7.

7. Conclusion

La détection automatique des polypes est une tâche très importante qui aide le praticien à établir un diagnostic appuyé. Ce chapitre a été consacré à l'implémentation de la solution basée sur l'utilisation du modèle YOLOv8 qui est, à ce jour, le dernier né de la série YOLO. A notre connaissance, c'est le premier travail concernant l'application de YOLOv8 dans la détection et la localisation des polypes. Les résultats obtenus sont très satisfaisants (mAP_50=96.93%, Precision= 96%, Recall= 92.92%, F1= 94.43%) qui dépasse les résultats obtenus avec YOLOv7.

Conclusion générale

L'utilisation de l'IA dans les systèmes de diagnostic assisté par ordinateur (CADx) équipés des algorithmes d'apprentissage automatique appropriés, et en particulier des méthodes d'apprentissage en profondeur, permet d'analyser de grandes quantités de données médicales pour aider les médecins à obtenir une détection très pertinente des anomalies en médecine.

Notre travail a porté sur la détection et la localisation automatisées des polypes lors d'examens de coloscopie, dans le but de fournir une aide à la décision pour une détection plus précise et précoce des polypes colorectaux et prévenir le cancer colorectal. Pour réaliser ce système nous utilisons une approche qui repose sur les réseaux de neurones convolutifs, en particulier le modèle YOLOv8. L'approche proposée montre des résultats très pertinents dans le domaine de la détection et localisation des polypes.

Pour l'entraînement de notre modèle, nous avons combiné deux ensembles de données qui sont Kvasir_SEG et BKAI-IGH NeoPolyp-Small et nous avons utilisé des techniques d'augmentation de données pour améliorer la performance de notre modèle. Ceci nous a permis d'obtenir des résultats très pertinents dans le domaine de la détection et la localisation des polypes.

Sur la base des résultats obtenus, le modèle YOLOv8 est une amélioration significative par rapport aux versions précédentes des modèles YOLO et donne de bons résultats par rapport à d'autres modèles de détection d'objets populaires en termes de précision et de vitesse.

Pour conclure, les principaux objectifs fixés dans ce travail ont été atteints avec succès. Cependant, il est important de souligner que des perspectives d'amélioration telles que :

- Continuer à améliorer la vitesse et la précision du modèle YOLOv8 pour la détection et la localisation automatisées des polypes en coloscopie soit en optimisant les algorithmes ou en augmentant la taille des data sets par exemple.
- Mettre en place une solution pratique pour l'utilisation du modèle dans le domaine médical, en prenant en compte les besoins et les contraintes spécifiques de ce domaine.
- Intégrer une fonctionnalité de classification des polypes pour fournir des informations plus précises sur leur type et leur niveau de risque.
- Collecter davantage de données pour améliorer la capacité du modèle à détecter les polypes dans des situations diverses.

Références

- [1] : Yacoub, R. B. (2019). Analogie et corps humain : cas du rapport analogique avec le tube digestif humain. *Revue Internationale d'Études en Langues Modernes Appliquées*, (12), 54-61.
- [2] : Kohler, C. (2011). L'appareil digestif. *Université Médicale Virtuelle Francophone*, 4-18.
- [3] : Wahab, H., Mehmood, I., Ugail, H., Sangaiah, A. K., & Muhammad, K. (2023). Machine learning based small bowel video capsule endoscopy analysis: Challenges and opportunities. *Future Generation Computer Systems*.
- [4] : Tolo, N., Apeti, S., Yves, B., Abdramane, T., Keita, K., Cissoko, M., ... & Cisse, S. O. (2023). La Fibroscopie Œsogastroduodénale chez le Sujet Âgé à Bamako: Une Étude de 121 Cas. *HEALTH SCIENCES AND DISEASE*, 24(6).
- [5] : Bernardini, D., Chaussade, S., Lapuelle, J., & Robaszkiewicz, M. La coloscopie.
- [6] : Haanstra, J. (2023). Gastrointestinal surveillance and management in Lynch syndrome.
- [7] : IMAD Nantes. (2020). La vidéo capsule endoscopique. Page consultée le 20 mai 2023, de <https://www.imad-nantes.org/la-video-capsule-endoscopique>
- [8] : Pourcelot, L., & Bouakaz, A. (2022). Développement de l'échographie et des techniques Doppler à l'Université de Tours. *Bulletin de l'Académie Nationale de Médecine*.
- [9] : Angermann, Q. (2018). 'Détection temps réel assistée par ordinateur pour le dépistage précoce du cancer colorectal en vidéocoloscopie' (Doctoral dissertation, Cergy-Pontoise).
- [10] : Johns Hopkins Medicine. (2020). Sporadic (Nonhereditary) Colorectal Cancer: Introduction.
- [11] : Karaman, A., Pacal, I., Basturk, A., Akay, B., Nalbantoglu, U., Coskun, S., ... & Karaboga, D. (2023). Robust real-time polyp detection system design based on YOLO algorithms by optimizing activation functions and hyper-parameters with artificial bee colony (ABC). *Expert Systems with Applications*, 221, 119741.

- [12] : Deeply Learning. (2018). Fonctionnement du neurone artificiel. Accédé le 25 mai 2023, <https://deeplylearning.fr/cours-theoriques-deep-learning/fonctionnement-du-neurone-artificiel/>
- [13] : Byrne MF, et al. (2019). Intelligence artificielle en temps réel pour la détection des néoplasies colorectales dans un essai randomisé. *Gut*, 68(10), 1813-1819.
- [14]: Wang P, et al. (2020). Deep learning for identifying polyps in colonoscopy: A systematic review and meta-analysis. *Journal of Medical Systems*, 44(7), 133.
- [15]: Chen, X., Zhang, K., Lin, S., Dai, K. F., & Yun, Y. (2021). Single shot multibox detector automatic polyp detection network based on gastrointestinal endoscopic Images. *Computational and Mathematical Methods in Medicine*, 2021.
- [16]: Wan, J., Chen, B., & Yu, Y. (2021). Polyp detection from colorectum images by using attentive YOLOv5. *Diagnostics*, 11(12), 2264.
- [17]: Quan, S. Y., Wei, M. T., Lee, J., Mohi-Ud-Din, R., Mostaghim, R., Sachdev, R., ... & Friedland, S. (2022). Clinical evaluation of a real-time artificial intelligence-based polyp detection system: a US multi-center pilot study. *Scientific Reports*, 12(1), 6598.
- [18] : Lui, T. K. L., Guo, C. G., & Leung, W. K. (2020). Artificial intelligence and colonoscopy: Current status and future perspectives. *Journal of Gastroenterology and Hepatology*, 35(5).
- [19]: Byrne, M. F., Chapados, N., Soudan, F., et al. (2019). Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*, 68(1), 94-100.
- [20]: Pohl, H., Figueiredo, J. C., & Rex, D. K. (2015). Study on missed lesions in colonoscopy: How many intervals should be evaluated? *Gastrointestinal Endoscopy Clinics of North America*, 25(2), 249-266.
- [21]: Rex, D. K. (2012). Challenges in the colonoscopic detection of nonpolypoid colorectal neoplasia. *Gastroenterology*, 142(2), 420-423.
- [22]: Ethical and Professional Considerations in Artificial Intelligence-Enabled Colonoscopy. (2020). *Gastrointestinal Endoscopy*, 91(5), 1222-1227.
- [23] : Mori, Y., Kudo, S. E., Misawa, M., et al. (2020). Artificial intelligence for colonoscopy. *Endoscopy*, 52(08), 716-720. doi:10.1055/a-1195-5804

- [24] : Byeon, S., Navarro-Alvarez, N., Hassan, C., et al. (2020). Artificial Intelligence in Gastrointestinal Endoscopy: The Future Is Almost Here. *Gut and Liver*, 14(6), 689-701. doi:10.5009/gnl19541
- [25]: Klare, P., Sander, C., Prinzen, M., Haller, B., Nowack, S., Abdelhafez, M., et al. (2019). Automated polyp detection in the colorectum: A prospective study (with videos). *Gastrointestinal Endoscopy*, 89(3), 576-582. doi: 10.1016/j.gie.2018.10.02
- [26] : Saba, D., Sahli, Y., Maouedj, R., Hadidi, A., & Medjahed, M. B. (2021). Towards artificial intelligence : concepts, applications, and innovations. *Enabling AI Applications in Data Science*, 103-146. doi: [10.1007/978-3-030-52067-0_6](https://doi.org/10.1007/978-3-030-52067-0_6)
- [27]: MESBAH, F. (2021). Détection d'objets par Deep Neural Network à l'aide du modèle YOLO en temps réel.
- [28]: Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9, 381-386.
- [29]: Kaur Gill, J. (2018). Automatic Log Analysis using Deep Learning and AI. Accès <https://www.xenonstack.com/blog/log-analytics-deep-machine-learning>.
- [30]: Torre, F. (1999). GloBo: un algorithme stochastique pour l'apprentissage supervisé et non-supervisé. In *Actes de la Première Conférence d'Apprentissage* (pp. 161-168).
- [31]: Guillet, V. (2022). Distillation de réseaux de neurones pour la généralisation et le transfert en apprentissage par renforcement (Doctoral dissertation, Toulouse, ISAE).
- [32]: Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1), e4150.
- [33]: Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7(3-4), 197-387.
- [34] : Thabet, I. (2021). Recalage des images médicales par apprentissage profond
- [35]: Boillet, M. (2023). Détection d'objets dans les documents numérisés par réseaux de neurones profonds (Doctoral dissertation, Université de Rouen).

Référence

- [36]: Mishra, M. (2020). Convolutional neural networks explained. Medium. (8. 2020.),
adresa:<https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939> (pogledano 31. 8. 2022.).
- [37]: Ikram, N. (2022). implémentation des Détecteurs de contours à Base de réseau Neurones Convolutifs CNN (Doctoral dissertation, faculté des sciences et de la technologie univ bba).
- [38]: Wikimedia Commons. (2022). File:typical cnn.png. [Capture d'écran]. Récupéré de https://commons.wikimedia.org/wiki/File:Typical_cnn.png
- [39]: BILAL, L. A. MÉMOIRE DE MASTERarchives. univ-biskra. Dz/.../14604/1/louam_abdelhak_bilal.pdf. 2019. 12. 16. · LOUAM ABDELHAK BILAL Le : Samedi 6 juillet 2019 Deep Learning basé sur les méthodes.
- [40]: O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.
- [41]: BENREKIA, M., MEKHTICHE, O., & BENGHERBIA, B. (2022). Etude et implémentation d'un système intelligent d'aide à la diagnostic automatique des défauts de roulements dans les machines tournantes.
- [42]: DataScientest().Convolutional Neural Network .Accédé le 1 mai 2023
<https://datascientest.com/convolutional-neural-network>
- [43]: Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. Neurocomputing, 187, 27-48.
- [44]: Patra, A. (2022). Deep Learning for Automated Polyp Detection and Localization in Colonoscopy (Master's thesis, OsloMet-storbyuniversitetet).
- [45]: Computer vision tasks [Capture d'écran]. Accédé le 15 mai 2023
https://www.researchgate.net/figure/Computer-vision-tasks-Adapted-from_fig5_346091812
- [46]: Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection.
- [47]: Brownlee, J. (2019). A gentle introduction to object recognition with deep learning.
- [48]: Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation.
- [49]: Girshick, R. (2015). Fast R-CNN.

- [50]: Ren, S., He, K., Girshick, R. B., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. CoRR, abs/1506.01497.
- [51]: Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Dans Conférence NIPS.
- [52] : Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. Procedia Computer Science, 199, 1066-1073.
- [53]: Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You only look once: Unified, real-time object detection. Preprint sur arXiv:1506.02640.
- [54]: Mekhennene, A., Sekhri, M., & Berkane, M. (2022). La Détection et le suivi des objets.
- [55] : Boillet, M. (2023). Détection d'objets dans les documents numérisés par réseaux de neurones profonds (Doctoral dissertation, Université de Rouen).
- [56]: Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 577-586).
- [57]: Terven, J., & Cordova-Esparza, D. (2023). A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. arXiv preprint arXiv:2304.00501.
- [58]: Liu, G., Nouaze, J. C., Touko Mbouembe, P. L., & Kim, J. H. (2020). YOLO-tomato: A robust algorithm for tomato detection based on YOLOv3. Sensors, 20(7), 2145.
- [59] : Chen, Z., Wu, R., Lin, Y., Li, C., Chen, S., Yuan, Z., ... & Zou, X. (2022). Plant disease recognition model based on improved YOLOv5. Agronomy, 12(2), 365.
- [60]: Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., ... & Wei, X. (2022). YOLOv6: A single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976.
- [61]: Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7464-7475).
- [62]: Łysakowski, M., Żywanowski, K., Banaszczyk, A., Nowicki, M. R., Skrzypczyński, P., & Tadeja, S. K. (2023). Real-Time Onboard Object Detection for Augmented Reality: Enhancing Head-Mounted Display with YOLOv8. arXiv preprint arXiv:2306.03537

[63]: Ang, G. J. N., Goil, A. K., Chan, H., Lee, X. C., Mustaffa, R. B. A., Jason, T., ... & Shen, B. (2023). A novel application for real-time arrhythmia detection using YOLOv8. arXiv preprint arXiv:2305.16727.

[64]: Ju, R. Y., & Cai, W. (2023). Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm. arXiv preprint arXiv:2304.05071.

[65]: Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., & Johansen, H. D. (2020). Kvasir-seg: A segmented polyp dataset. In MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26 (pp. 451-462). Springer International Publishing.

[66]: Hanoi University of Science BK.AI, Technology incorporation with Institute of Gastroenterology, and Hepatology (IGH) Vietnam. Bkaiigh neopolyp a dataset for colonoscopy polyp segmentation and neoplasm characterization.

[67]: Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft coco: Common objects in context, 2014.

[68]: <https://app.roboflow.com> Accédé le 22/06/2023.

[79]: Google Colab tutorial - Great Learning (mygreatlearning.com)
<https://www.mygreatlearning.com/blog/google-colab-tutorial/> Accédé le 07/06/2023

[70]: Python Programming Language GeeksforGeeks <https://www.geeksforgeeks.org/python-programming-language> Accédé le 07/09/2021

[71]: <https://roboflow.com/> Accédé le 07/06/2023

[72]: About - OpenCV <https://opencv.org/about/> Accédé le 07/06/2023

Référence

[73] : <https://www.tresfacile.net/la-bibliotheque-gui-python-tkinter> Accédé le 07/06/2023

[74] : PyCharm. (2023, février 2). Wikipédia, l'encyclopédie libre. Page consultée le 11:14, février 2, 2023 à partir

<http://fr.wikipedia.org/w/index.php?title=PyCharm&oldid=201014062>:

[75] : Asyraf, M. S., Isa, I. S., Marzuki, M. I. F., Sulaiman, S. N., & Hung, C. C. (2021). CNN-based YOLOv3 comparison for underwater object detection. Journal of Electrical and Electronic Systems Research (JEESR), 18, 30-37.