

*The brave new world  
of artificial  
intelligence for  
digestive endoscopy*

Xavier Dray<sup>1,2</sup>  
Romain Leenhardt<sup>1</sup>  
Aymeric Histace<sup>2</sup>  
Aymeric Becq<sup>1</sup>

<sup>1</sup> Sorbonne Université & APHP, Hôpital  
Saint Antoine, 184, rue du Faubourg  
Saint-Antoine, 75012 Paris

<sup>2</sup> ETIS, Université Paris-Seine, Université  
de Cergy-Pontoise, ENSEA, CNRS, Cergy-  
Pontoise

@ Correspondance : X. Dray  
xavier.dray@aphp.fr

## Intelligence artificielle et endoscopie : le meilleur des mondes ?

### ▼ Résumé

L'intelligence artificielle (IA) a pour objet de simuler l'intelligence humaine. Il s'agit d'une science cognitive qui fait appel à la neurobiologie et à la logique (résolution des problèmes, apprentissage profond, réseaux neuronaux), à l'informatique (puissance de calculs des ordinateurs, internet), et qui se nourrit de bases de données. Elle trouve ses premières applications en santé – et ses premières autorisations de mise sur le marché – dans l'exploitation des mégadonnées (épidémiologie, médecine prédictive) et dans l'analyse des signaux (tracés ECG et EEG, imagerie médicale, anatomopathologie, dermatologie, ophtalmologie...). Leur champ d'application en endoscopie digestive est également très vaste, englobant le dépistage, le diagnostic, la caractérisation, la thérapeutique, le pronostic, et cela pour tout type d'intervention. Le nombre de publications dans ce domaine augmente de façon exponentielle. Les premières réalisations concernant la détection et la caractérisation de polypes coliques assistées par ordinateur devraient faire prochainement l'objet d'une industrialisation. La pré-lecture de vidéocapsules endoscopiques par des systèmes d'apprentissage profond, en réseau, est également un modèle très illustratif du développement de l'IA en endoscopie. Cette révolution technologique est susceptible d'améliorer les performances des médecins et donc la qualité des soins. Elle mérite d'être accompagnée par les endoscopistes, non seulement concernant ses aspects techniques et cliniques, mais également pour répondre aux nouvelles questions qui en découlent, notamment sur les places relatives du médecin (moins technicien, plus humain ?) et de la machine (assistante ou autonome ?), sur la responsabilité médicale (homme ou machine ?), sur le remboursement des actes (médecin ou prestataire ?).

• **Mots clés** : intelligence artificielle, mégadonnées, apprentissage profond, réseaux neuronaux

### ▼ Abstract

*Artificial intelligence (AI) aims to simulate the human intelligence. It is a cognitive science which relies on neurobiology, logical and critical thinking (problem solving, deep learning, neural networks), computing sciences (calculation, internet), and on databases. Big data exploitation (epidemiology, predictive medicine) and "signals" analysis (EKG, EEG, imaging, pathology, dermatology, ophthalmology...) were the first successful application of AI in healthcare, followed by government approval. AI has a vast spectrum of potential applications in digestive endoscopy as well. AI can be used for screening, diagnosis, characterization, treatment, and prognosis evaluation, in a wide array of procedures. The quantity of published work in this field is thriving. Computer-assisted detection and characterization of colonic polyps for instance, were amongst the first successful applications of AI, and should be commercially available shortly. The automated reading of a capsule*

Pour citer cet article : Dray X, Leenhardt R, Histace A, Becq A. Intelligence artificielle et endoscopie : le meilleur des mondes ? Hépatogastro et Oncologie Digestive 2019 ; 26 : 319-331. doi : 10.1684/hpg.2019.1754

endoscopy, based on a network of machine learning systems, is also very demonstrative of what AI will be able to accomplish in the next future. It is believed that AI will significantly improve diagnostic performances and thus the quality of care. Today, endoscopists should not only promote this technological revolution, but also address new issues in the field of AI, regarding the respective roles of physicians (focused on ethics and patient-relations) and AI-machines (assistants vs autonomous), as well as responsibility (physicians vs. manufacturing companies), and reimbursement (physician vs manufacturing companies).

• **Key words:** artificial intelligence, big data, deep learning, neural networks

## Introduction

Après la finance, le marketing, les jeux (go, poker, échecs...), les véhicules autonomes, et tant d'autres initiatives, l'intelligence artificielle (IA) s'est invitée dans le monde de la santé et de la médecine. Pour ses défenseurs, elle est la promesse de choix, diagnostics et thérapeutiques, plus précis et rapides, de capacités de recherche démultipliées par le big data, outrepassant les capacités de travail et de création des meilleurs d'entre nous, sans enjeu d'affect, d'ego, de salaire ou de récupération du temps de travail. Pour ses détracteurs, l'IA est un danger de dépersonnalisation de la médecine, de perte des fondamentaux cliniques. Qu'en sera-t-il demain si une machine décide de soins qui déterminent notre bien-être physique et psychique (une définition de la santé) et parfois même notre survie ? Sans doute y a-t-il un juste milieu. Comme pour tous les outils inventés par l'Homme (du feu jusqu'à l'énergie nucléaire, en passant par le couteau), il y a de bons et mauvais usages. La connaissance de ces outils contribue certainement à leur meilleure maîtrise (technique mais aussi éthique) par leurs utilisateurs et bénéficiaires espérés de demain : nous-mêmes. L'objet de cette revue est de faire le point sur l'avancement de l'IA appliquée en médecine en général, et en endoscopie en particulier, en insistant donc sur le traitement des signaux et de l'image.

## Qu'est-ce que l'intelligence artificielle ? Et comment ça marche ?

L'intelligence a de multiples définitions. Nous retiendrons deux d'entre elles, issues du dictionnaire Larousse : (a) « l'ensemble des fonctions mentales ayant pour objet la connaissance conceptuelle et rationnelle » et (b) « l'aptitude (...) à s'adapter à une situation, à choisir des moyens d'action en fonction des circonstances ». Une définition de l'IA est, selon ce même dictionnaire, « un ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence ».

**/// Faire faire par une machine des tâches que l'homme accomplit en utilisant sa propre intelligence ///**

L'IA est une création humaine, et peut donc être perçue comme un prolongement, une optimisation, un renfor-

cement, de l'intelligence humaine, du moins tant qu'elle reste sous contrôle et supervision d'une intelligence humaine. L'IA fait partie des sciences cognitives et fait elle-même appel à d'autres sciences.

**/// L'intelligence artificielle n'est plus une fiction, elle est une science ///**

L'IA s'inspire d'abord largement de la *neurobiologie*. Le terme de « réseaux neurones artificiels » (un des outils de l'IA) y fait largement écho. On considère l'œil comme un appendice du cerveau, comme un organe sensoriel qui capte environ 70 % des informations transmises au cerveau, sous forme de signal électrique via le nerf optique. Nous verrons que le traitement informatique de l'image (un « signal », une information, parmi d'autres) tente de reproduire ce couplage entre œil et cerveau.

**/// L'intelligence artificielle est à la croisée de la neurobiologie, de la logique, et des sciences computationnelles ///**

L'IA fait également appel à l'*apprentissage machine* (ou *machine learning*), à travers des méthodes de résolution des problèmes appelées « algorithmes ». Pour certaines de ces approches, on peut présenter ces algorithmes comme « artisanaux » (*handcrafted methods*). L'homme fournit à la machine les données d'apprentissage et la « vérité terrain » associée, décrivent au mieux les éléments caractéristiques à extraire, dit à la machine ce qu'elle doit apprendre, comment elle doit le faire, quelles décisions elle doit prendre en fonction des informations qu'on lui donne durant l'étape d'apprentissage. À titre d'illustration, la reconnaissance d'un visage au sein d'une série de photos par un logiciel d'analyse d'images, fait appel à ce type de technique « artisanale », « faite maison ». L'expert humain soumet à la machine une série d'images bien caractérisées et descriptibles (des visages) et des images sans cette image (d'autres contenus, sans visage). Les caractéristiques du contenu de l'image (couleur, luminosité, contraste, texture, points saillants...) sont traduites en vecteurs mathématiques de caractéristiques de plus ou moins grandes tailles. L'homme peut faire cette description pour la machine (éventuellement aidé par la machine). Avec un nombre plus ou moins important d'images, la machine saura reconnaître des visages, et continuera à apprendre à les

reconnaître de mieux en mieux si, au fur et à mesure de son utilisation, on continue à lui dire ses succès et ses erreurs (on parle alors d'apprentissage actif). Cette méthode a l'avantage d'être relativement aisée (peu coûteuse en ressources de calcul en particulier) mais elle est peu généralisable à une autre application : si votre logiciel de visionnage de photos vous permet de sélectionner des visages au milieu de vos photos de vacances, et même d'isoler le visage de votre belle-mère, il ne sera pas capable en revanche de sélectionner les photos contenant des chats ou des arbres sans réapprendre intégralement ce qu'est un chat ou un arbre au sens de vecteurs de caractéristiques.

### /// Des algorithmes « artisanaux » (handcrafted) peuvent être élaborés par l'homme ///

Alternativement, la machine elle-même peut élaborer des algorithmes et choisir les caractéristiques les plus pertinentes à l'application, que nous pourrions présenter comme des « automatismes », à travers des expériences répétées, ou à travers l'assimilation d'autres expériences à une nouvelle tâche à accomplir. C'est ce type d'algorithme, plus complexe, que nous mettons en œuvre pour des choix très élaborés. Ces algorithmes ne peuvent pas être dessinés avec une série de choix binaires (oui/non, succès/échec...). Les relations entre les niveaux de décision ont de multiples entrées et sorties, interrelations, et degré d'influence dans la prise de décision. On pourrait les comparer à des décisions que nous prenons sur la base de l'« expérience », du « savoir-faire »,

sans pour autant savoir facilement expliquer les raisons de notre choix si ce n'est par cette accumulation de nombreux exemples issus de la pratique. La reproduction artificielle de ce modèle d'intelligence fait appel à une forme différente d'algorithmes, issue de l'« apprentissage profond » ou *deep learning*. En vision par ordinateur, ces approches cherchent véritablement à reconstruire les mécaniques d'extraction de données faites par l'œil et se caractérisent par des vecteurs de caractéristiques pouvant être de dimension dépassant le million de paramètres. Notons que certaines techniques d'IA dans le traitement de l'image combinent méthodes artisanales et *deep learning*.

### /// En apprentissage profond (deep learning), la machine forge sa propre « expérience », élabore ses propres algorithmes ///

Attardons-nous quelques lignes sur les réseaux de neurones convolutifs (*neural networks*) (figure 1), qui sont des outils aujourd'hui indispensables dans la panoplie des *machine learners*. Ces réseaux de neurones « profonds » utilisent un nombre croissant de « couches » (aussi appelées *layers*) ou « filtres » de décisions (utilisant très majoritairement des convolutions locales des informations images, comme l'œil, d'où leur nom). À l'inverse des « *handcrafted features* », ces réseaux analysent l'ensemble des caractéristiques de l'image par le passage successif au travers des couches qui le constituent (contours, homogénéité, orientation, etc.). Le type, le nombre et l'ordre de ces filtres impactent les performances

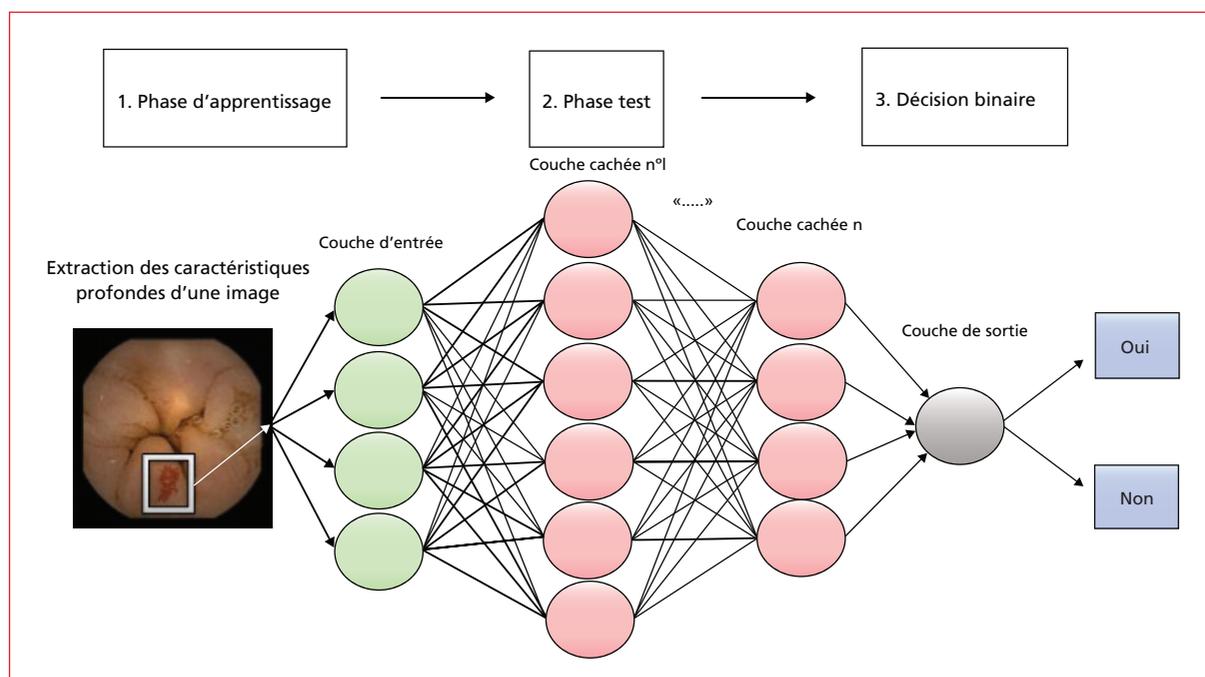


Figure 1 • Structure de réseaux neuronaux.

de l'algorithme. Durant la phase d'apprentissage, différentes combinaisons (type, nombre, ordre) sont testées par la machine, et la meilleure combinaison est retenue pour être appliquée. Le principal avantage de cette technique est que l'apprentissage par la machine est très généralisable : un tel processus ne se limitera pas à reconnaître un visage, mais pourra reconnaître d'autres « formes » caractérisées (chat, arbre...) sans pour autant modifier intégralement la structure du réseau. Ses principaux inconvénients en regard sont que cet apprentissage repose sur des bases de données très larges, et que les calculs peuvent être complexes et longs. Enfin, il est difficile, sinon impossible, pour un humain – même expert en IA – de considérer « maîtriser » son outil. Il ne peut qu'en mesurer les performances, et les améliorer par des artifices techniques, mais les schémas de décision de ces algorithmes peuvent devenir impénétrables à l'esprit humain.

▄▄ **Les réseaux neuronaux sont une méthode très répandue de l'apprentissage profond** ▄▄

Dans l'application de ces principes de *machine learning* (qu'ils soient de type *handcrafted* ou *deep learning*) à

l'analyse de l'image, on fera quasiment systématiquement appel au traitement statistique des données. La technique de traitement statistique la plus connue et la plus utilisée est l'*analyse en composantes principales* (ACP) où sont étudiées des combinaisons et corrélations de variables (*figure 2*). L'analyse en composantes principales ne relève pas de la recherche de combinaisons binaires (par exemple blanc/noir), mais de la recherche de combinaisons multifactorielles (blanc-grand-irrégulier... vs. noir-petit-régulier...) dans des espaces à plusieurs dimensions qui résument l'information que contiennent les images.

Enfin, l'IA fait appel à l'*informatique*. Le traitement des données (dans notre propos, l'image) nécessite des puissances de calculs importantes. Ainsi, l'approche des réseaux neuronaux convolutifs, proposée à la fin des années 1960, était alors peu applicable de par les ressources limitées des microprocesseurs. Les immenses progrès réalisés dans les technologies informatiques (parallélisations des calculs par exemple), combinés à la mémoire démultipliée et à la puissance de calcul qu'offrent aujourd'hui les processeurs, ont rendu possible l'avènement de l'IA. Le caractère interconnecté de nos outils

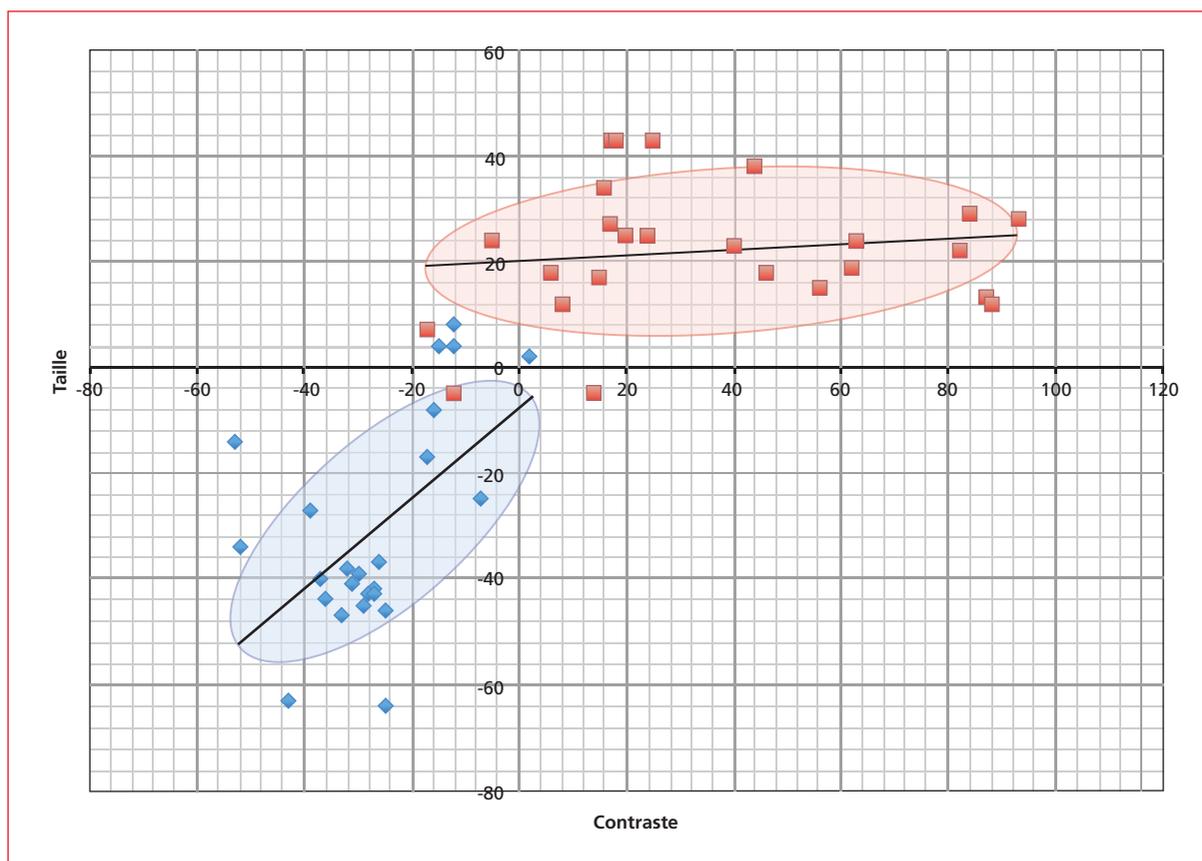


Figure 2 • Analyse en composantes principales. L'exemple donné ici illustre de façon volontairement caricaturale deux groupes d'images nettement distinctes selon deux de leurs caractéristiques (ici, leur texture et leur contraste), c'est-à-dire en deux dimensions. L'analyse s'applique en fait dans des espaces à n dimensions qui résument les informations que contiennent des séries images préalablement transformées en matrices mathématiques.

informatiques (*l'Internet of Things*) permet de plus d'accumuler des masses de données suffisantes pour parachever l'apprentissage profond par les machines, et de l'améliorer encore alors que les données s'accumulent continuellement.

**/// L'augmentation exponentielle des puissances de calcul des ordinateurs, et leur mise en réseau par internet, ont rendu possible l'avènement de l'intelligence artificielle ///**

## Quelles sont les principales avancées en médecine ?

### Le champ infini du big data

Un pan d'application gigantesque de l'IA concernera demain la médecine prédictive, via le big data, en « fouillant » dans les données (pour peu qu'elles soient interconnectées) de dossiers médicaux, de recours aux soins, de consommations de médicaments (achat et observance), des données plus générales (socio-économiques, géographiques...) et demain peut-être des données biologiques (voire génétiques) pour déterminer si un patient cumule des facteurs de risque et/ou des maladies. Au-delà de la démarche individuelle, pour le patient, les machines pourront travailler sans relâche à la recherche épidémiologique. De nombreuses institutions médicales, systèmes d'assurance, startups et géants de l'informatique, développent de telles stratégies et des « entrepôts de données » qui pourraient, demain, nourrir cette ambition. De nombreux projets voient par ailleurs le jour dans l'aide à la rédaction de compte-rendu, l'aide à la décision thérapeutique, l'aide à la prescription, le suivi thérapeutique, etc. La plupart des acteurs de ce secteur devinent que les principaux écueils à ce développement ne sont pas techniques. Les outils de l'IA sont certainement perfectibles, mais déjà bien présents, et très performants. La première difficulté à surmonter est de leur fournir des données de bonne qualité, et à traiter dans des quantités astronomiques.

**/// Les données sont la « matière première » de l'intelligence artificielle ///**

Pour autant, ces données médicales sont extrêmement sensibles, confidentielles, et des garde-fous éthiques et juridiques sont nécessaires avant de les interconnecter avec d'autres données pertinentes, éventuellement également sensibles (état civil, revenus, par exemple), et de les diffuser à des acteurs non-institutionnels (assurances, géants de l'informatique, etc.). L'interprétation de ces données soulève également des questions nouvelles. Traiter autant de variables (des centaines, des milliers...) à des échelles jusque-là insoupçonnées (des millions, dizaines de millions, et plus) d'individus

(pas nécessairement malades), appellera de nouvelles approches statistiques. Des associations statistiques mettront certainement en évidence de nombreux liens entre des variables, sans que pour autant elles soient toutes « cliniquement pertinentes » ou qu'il y ait des relations de cause à effet, et encore moins que le sens des relations (de cause à effet) soit évident. La « logique » du *machine learning* étant souvent insondable, une science de l'interprétation scientifique et médicale de ses résultats devra naître.

**/// Les mégadonnées (big data) de santé, issues de l'ère numérique, dépassent les capacités humaines d'analyse, mais sont exploitables par l'intelligence artificielle ///**

### L'analyse du signal, une aide au diagnostic en plein essor

Les applications de l'IA les plus immédiates en médecine concernent le traitement du signal, et notamment de l'image. Quelques illustrations récentes et exemplaires en sont livrées ici, pour montrer ce qui pourrait être réalisé demain en hépato-gastroentérologie, et notamment en endoscopie digestive.

Une des applications les plus abouties à ce jour est une plateforme en ligne, française, d'interprétation d'ECG ambulatoires. La technologie repose sur l'internet et sur un réseau de neurones qui a été entraîné sur plus de 500 000 enregistrements ECG. Des données préliminaires montrent une sensibilité de 96,9 % et une spécificité de 98,5 % pour le diagnostic de fibrillation auriculaire [1]. La technologie a reçu une autorisation de la *Food and Drug Administration* (FDA) pour l'assistance dans la reconnaissance de 10 troubles du rythme cardiaque aux États-Unis d'Amérique, et en Europe, un marquage CE pour la détection de plus de 100 anomalies. Des approches similaires sont en développement pour l'aide à l'interprétation, plus complexe encore, des EEG. Une autre des applications les plus abouties est le dépistage automatisé de la rétinopathie diabétique [2]. Entraîné sur une base de près de 130 000 fonds d'œil, l'outil basé sur un réseau de neurones avait une sensibilité entre 87 % et 97 % et une spécificité entre 93,9 % et 98,5 %. Ce système a été la première solution d'IA approuvée par la FDA.

**/// Les systèmes d'intelligence artificielle reçoivent leurs premières autorisations de mise sur le marché ///**

Les applications en santé de l'IA dévolues à la dermatologie ont eu un écho important dans les médias. Un outil a appris à différencier plus de 2 000 maladies de la peau, à partir de 129 450 images prises sur le web, avec un réseau de neurones lui aussi disponible sur le net. La machine

distinguaient carcinomes spinocellulaires et kératoses bénignes d'une part, mélanomes malins et nævi bénins d'autre part, avec des précisions diagnostiques de 96 % et 90 % des cas, respectivement, tandis les dermatologues le faisaient dans 95 % et 76 % des cas [3]. Depuis, un autre travail a montré que, dans le diagnostic dermatoscopique de mélanome, la précision diagnostique d'un algorithme (86 %) était significativement plus élevée que celle d'un groupe international de 58 dermatologues (79 %), incluant 30 experts [4].

### /// Certains systèmes d'intelligence artificielle dépassent les performances de médecins experts ///

C'est sans doute en imagerie radiologique que l'IA imprime le plus sa marque sur la médecine actuellement [5]. Tous les champs de la discipline sont approchés (radiographie, échographie, scanner, IRM), dans toutes ses applications (y compris digestives). Ainsi, des machines ont des performances diagnostiques supérieures à 90 % (qui surpassent parfois celles des urgentistes ou des radiologues eux-mêmes) par exemple à partir de radiographies de thorax pour les diagnostics de tuberculose, ou de pneumonie, mais aussi dans le dépistage, avec des initiatives aux résultats encourageants concernant par exemple le cancer du sein ou le cancer du poumon par scanner faiblement irradiant. Les grands acteurs de l'industrie de l'imagerie médicale, mais aussi les géants de l'informatique, s'intéressent de près à ce développement et y investissent des moyens colossaux. La Société Française de Radiologie développe un système dédié à l'imagerie médicale, mettant à profit les PACS (*Picture Archiving and Communication Systems*, système d'archivage et de partage des images) pour alimenter d'immenses bases de données d'apprentissage pour les systèmes d'IA de demain. L'anatomopathologie, dont les fondamentaux reposent sur l'analyse d'images, a déjà connu récemment deux révolutions : celle de l'immunohistochimie, puis celle de la génomique. La troisième – l'IA – est en marche, mais ces développements restent pour l'instant modestes.

### /// Les métiers de l'imagerie médicale vont être bousculés par l'intelligence artificielle ///

Ces avancées sont souvent liées à l'organisation de *challenges* internationaux par des géants de l'informatique, des industriels, souvent en partenariat avec des sociétés savantes. Des *databases* contenant des milliers d'images pathologiques et d'images contrôles sont fournies à des équipes de recherches (elles-mêmes quelques centaines ou milliers). Les chercheurs sont évalués sur les performances de leurs outils sur différentes tâches (*Dream challenge* par exemple pour le dépistage du cancer du sein, *Annual Data Science Bowl* pour le dépistage du cancer du poumon...). Les équipes de

recherche (institutionnelles, startups, ou autres) valorisent considérablement leur travail (médiatisation, brevets, partenariat industriel) à l'issue de ces compétitions.

### /// Des défis (challenges) internationaux pour promouvoir la recherche ///

## Premières applications en endoscopie digestive, et développements attendus

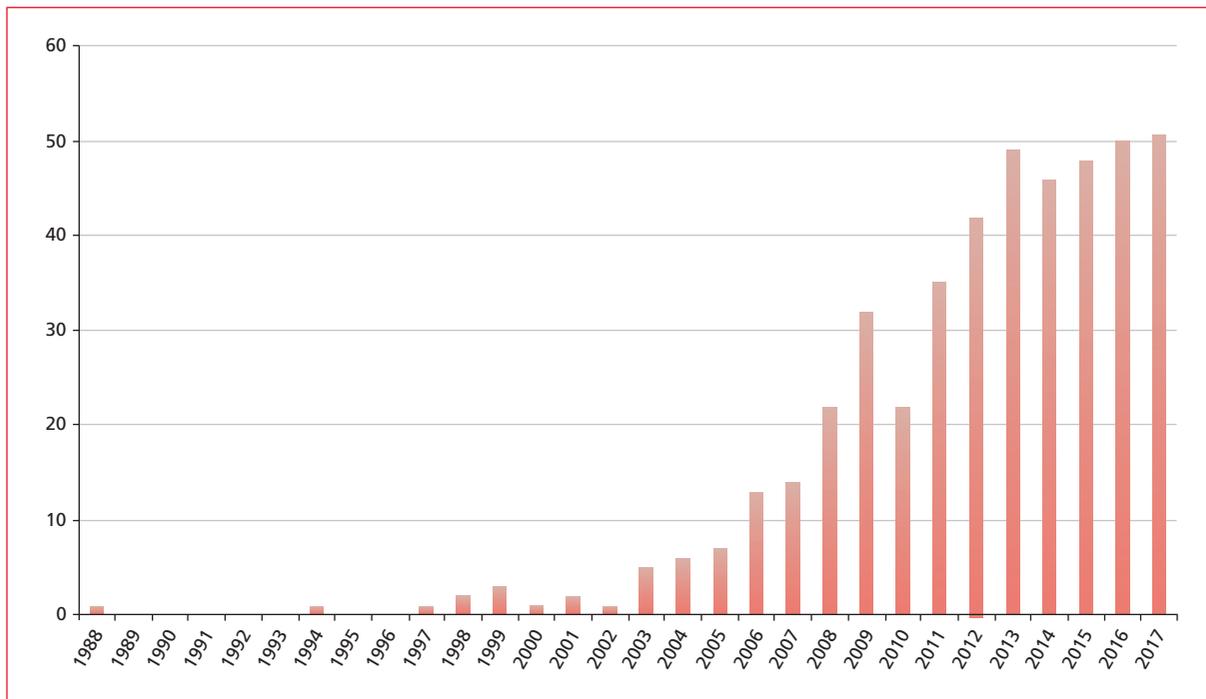
Le nombre de publications qui se réfèrent à l'IA en endoscopie digestive est en croissance exponentielle (*figure 3*), et toutes ne peuvent être détaillées ici. Cette recherche s'intéresse à tous les aspects de l'endoscopie : progression des dispositifs, qualité de la préparation, détection des lésions, caractérisation des lésions, pronostic. Nous illustrons ici des applications en voie de développement d'une part en coloscopie (parce que cette procédure est emblématique de l'activité des endoscopistes et que l'IA répond effectivement à des besoins non couverts), et d'autre part en vidéocapsule endoscopique du grêle (parce qu'il s'agit d'un modèle de développement de systèmes autonomes).

### /// L'endoscopie digestive n'échappe pas à l'élan nouveau que l'intelligence artificielle donne à la médecine moderne ///

## L'exemple emblématique du développement de l'intelligence artificielle en coloscopie

### • Assistance à l'examen de la surface muqueuse en coloscopie

Des systèmes d'aide à la progression du coloscope sont en cours de développement. Nous prenons le parti de détailler ici un système d'assistance au retrait, l'EM-Automated-RT (EndoMetric Inc, Ames, IA). Son objectif est de préciser en temps réel si une surface muqueuse suffisante a été examinée par l'opérateur [6, 7]. Le système examine les caractères net/flou et propre/sale de chaque image, ainsi que la paroi muqueuse. L'image endoscopique est divisée en cadrans, et le système valide si chaque cadran a bien été exposé et examiné, donnant un score à l'opérateur. L'évaluation clinique a été réalisée auprès de 10 internes en 3<sup>e</sup> année de spécialité, ayant réalisé 483 coloscopies, avec ou sans l'aide du système, et en comparaison de la relecture par des experts des vidéos correspondantes. L'utilisation du système améliorait significativement les moyennes des scores d'évaluation muqueuse, des scores de lavage de résidus, des scores de distension intestinale, et des temps de retrait [7].



**Figure 3** • Publications référencées, relatives à l'intelligence artificielle en endoscopie digestive, au cours des trente dernières années. D'après Kashin S. Artificial intelligence: the rise of the machines. United European Gastroenterology Week. 20-24 octobre 2018, Vienne, Autriche.

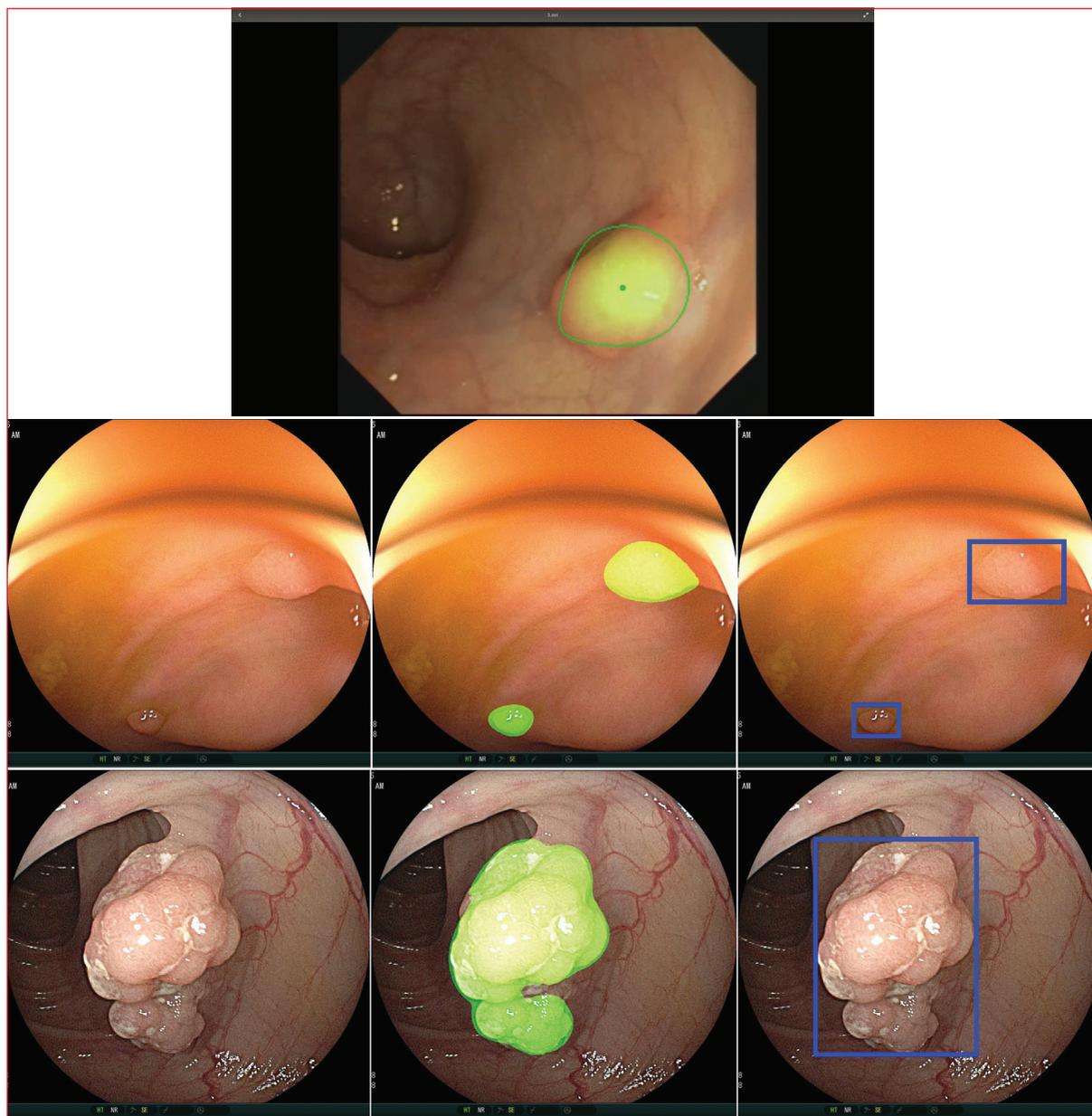
### /// L'opérateur guidé par un système d'intelligence artificielle améliore significativement la surface de côlon qu'il examine ///

#### • Aide à la détection de polypes en vidéocoloscopie

Plusieurs systèmes d'IA dédiés à la détection de polypes ont été développés (*figure 4*) [8, 9]. Certains sont arrivés à maturation et ont récemment été évalués cliniquement. Un des plus aboutis est basé sur un algorithme d'apprentissage profond à partir de 27 113 images de coloscopie (dont 5 541 avec polypes) issues de 1 290 patients [9]. La première phase de validation a essentiellement été réalisée à partir d'images fixes sur un autre lot indépendant d'images fixes de 1 138 autres patients. Les performances par image (et non par patient) étaient les suivantes : sensibilité 94,4 %, spécificité 95,9 %, précision diagnostique 98,4 %. Le temps de latence était pour chaque image de  $76,8 \pm 5,6$  ms, ce qui correspond à une cadence d'analyse de 25 images par seconde, compatible avec les cadences de capture d'image des processeurs d'endoscopie actuels (25 à 60 images/seconde). Ces résultats ont ouvert la voie à une évaluation clinique [10]. Dans un essai monocentrique, prospectif et randomisé, récent, 1 058 patients (porteurs de 767 polypes, 422 adénomes et 31 adénomes festonnés) étaient randomisés pour avoir leur coloscopie avec *versus*

sans l'aide de ce système de détection de polypes en temps réel (avec alerte visuelle et sonore). Le taux de détection de polypes était de 29,10 % dans le groupe sans détection assistée par ordinateur (DAO) vs. 45,02 % dans le groupe avec DAO (OR = 1,995,  $p < 0,001$ ). Le nombre moyen de polypes était de 0,50 dans le groupe sans DAO vs. 0,95 dans le groupe avec DAO ( $\times 1,89$ ,  $p < 0,001$ ). Le taux de détection d'adénomes étaient de 20,34 % dans le groupe sans DAO vs. 29,12 % dans le groupe avec DAO (OR = 1,61,  $p < 0,001$ ). Le nombre moyen d'adénomes était de 0,31 dans le groupe sans DAO vs. 0,53 dans le groupe avec DAO ( $\times 1,72$ ,  $p < 0,001$ ). Aucun polype n'était manqué par le DAO, et il indiquait seulement 39 faux positifs. Les analyses par sous-groupes montraient une amélioration significative des performances de l'opérateur pour toutes les localisations sauf le cæcum et le rectum, pour les polypes non pédiculés, pour les polypes de moins de 10 mm (et plus significativement encore pour les moins de 5 mm), et pour les polypes hyperplasiques et adénomateux (mais pas pour les festonnés). Des résultats similaires ont été obtenus par un autre système basé sur l'apprentissage machine, avec une sensibilité de 96,9 %, une spécificité de 95 % et une précision diagnostique de 96,4 % [11], apportant une validation externe à cette approche.

### /// Un essai contrôlé randomisé a démontré que l'intelligence artificielle augmentait le taux de détection d'adénomes ///



**Figure 4** • Illustrations de différents systèmes temps-réel de détection de polypes coliques assisté par ordinateur. Les images sont extraites d'un examen complet réalisé en conditions réelles. Les algorithmes issus du machine learning permettent de signaler la présence de polypes pendant l'examen. En haut : d'après notre prototype [6]. En bas : d'après [8, 9], avec l'aimable permission du Dr Tyler M Berzin (Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center, Harvard Medical, Boston, MA, États-Unis).

### **Aide à la caractérisation de polypes en vidéocoloscopie**

Les systèmes d'IA développés pour la caractérisation de polypes coliques viennent enrichir les options de chromoendoscopie virtuelle et de magnification actuellement proposées par les firmes d'endoscopie [12-15]. Les systèmes le plus aboutis aujourd'hui sont couplés au *Narrow Band Imaging* (NBI, Olympus, Tokyo, Japon). Le

couplage coloration-zoom-IA consolide encore le concept de « biopsie optique » en temps réel, aboutissant à des performances au-delà des critères de validation PIVI tels que définis notamment par l'American Society of Gastrointestinal Endoscopy (ASGE), avec une valeur prédictive négative de plus de 90 % pour la reconnaissance de polypes adénomateux diminutifs. Des images édifiantes d'endocystoscopie, par le système EndoBRAIN, en renfort du NBI, ont circulé quelques années dans les congrès.

Elles ont fasciné le public, montrant l'interprétation d'images en très fort grossissement, avec un niveau de probabilité de diagnostic hyperplasique vs. adénome, en 0,3 seconde en moyenne. Ces vidéos ont fait la preuve définitive du concept. Néanmoins, rares sont les unités équipées d'endocystoscopie, et, par ailleurs, les performances évaluées sur 100 lésions (avec preuve anatomopathologique à l'appui) étaient légèrement en dessous des seuils PIVI : sensibilité 85 %, spécificité 98 %, précision diagnostique 90 % [13]. L'apprentissage profond a amélioré ces performances, sans recours à l'endocystoscopie, et avec un temps de latence moyen (50 ms) compatible avec une utilisation en temps réel. Depuis, les résultats de l'entraînement d'un système de *deep learning* à partir de plus de 60 000 images sélectionnées issues de 223 vidéos, en NBI, de muqueuse normale et de polypes coliques en vision proche et très proche (technologie *near focus*, Olympus, Tokyo, Japon), ont été publiés [15]. Le système proposé affiche un degré de « crédibilité » de diagnostic de muqueuse normale, de polype hyperplasique (NICE 1) ou d'adénome (NICE 2) (figure 5). Sur un jeu indépendant de 125 séquences vidéos de polypes de 10 à 20 secondes, les performances du système étaient comparées à l'analyse anatomopathologique, et dépassaient les critères PIVI : sensibilité 98 %, spécificité 83 %, précision diagnostique 94 %. Ces résultats ouvrent la voie à l'évaluation clinique du système (à l'échelle de la vidéo, et donc du patient), avec sans doute l'intégration prochaine de la reconnaissance automatisée des adénomes festonnés [16], et dans un futur plus lointain le guidage de la démarche thérapeutique (délimitation des marges, choix de la technique [17]...) et du pronostic (proposition en temps réel de complément chirurgical, du calendrier de surveillance par exemple) [18, 19].

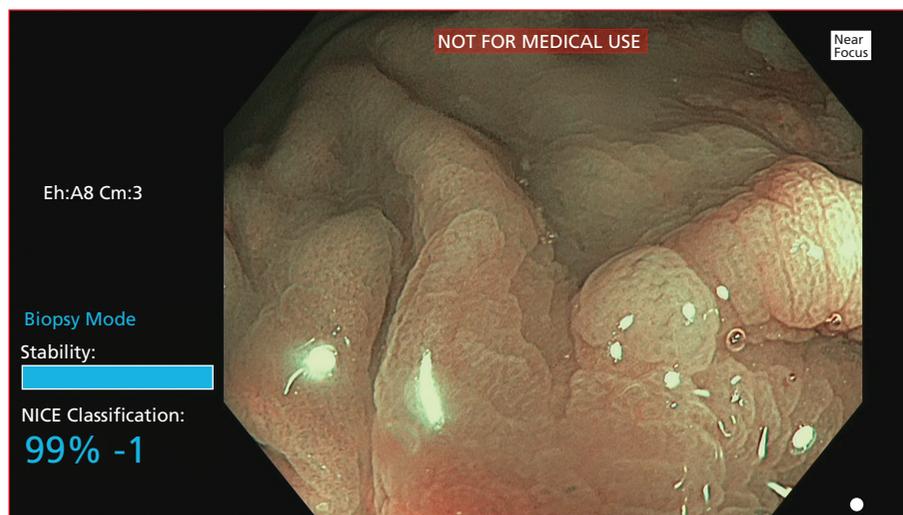
**/// La caractérisation de polypes assistée par ordinateur, en temps réel, est hautement performante ///**

### **Une illustration d'une recherche émergente : l'intelligence artificielle appliquée à la vidéocapsule endoscopique**

La vidéocapsule endoscopique (VCE) représente une faible part de l'activité en endoscopie digestive aujourd'hui, avec des enjeux de santé publique actuellement moindre, et des indications essentiellement d'exploration de l'intestin grêle. Le couplage de cette technologie issue de la miniaturisation avec l'IA représente un modèle de développement intéressant. Les images captées sont archivées et analysées en différé, ce qui permet de relativement s'affranchir des contraintes de l'analyse en temps réel.

#### **• Aide à l'évaluation de la qualité de la préparation en vidéocapsule endoscopique du grêle**

Les performances de la VCE peuvent être diminuées par une mauvaise visualisation de la muqueuse du fait du contenu intestinal (fluides, débris, bile), de la présence de bulles, ou en raison d'une luminosité trop faible, sans possibilité de lavage ou d'aspiration. La proportion d'examen par VCE mal préparés varie ainsi entre de 17 % et 32 % dans la littérature [20]. En l'absence d'échelle validée de la qualité de visualisation lors de VCE de l'intestin grêle, il n'existe pas de consensus quant aux modalités de la préparation intestinale pour optimiser



**Figure 5** • Reconnaissance d'un polype hyperplasique par un système temps-réel de caractérisation de polypes coliques assisté par ordinateur. Avec l'aimable permission du Dr Michael Byrne (Division de Gastroentérologie, Vancouver General Hospital et University of British Columbia, Canada).

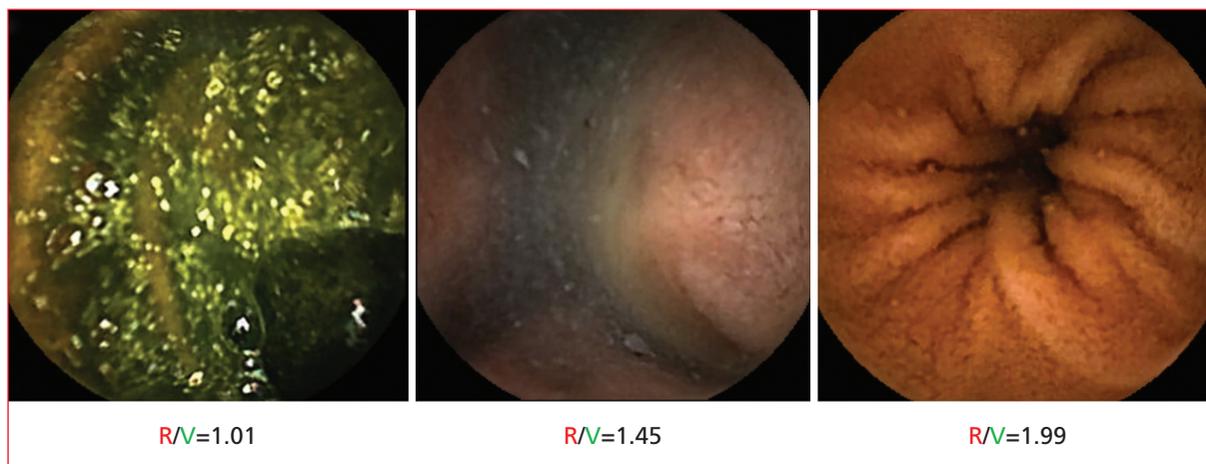


Figure 6 • Exemples de scores colorimétriques électroniques (ratio de pixels rouge/vert) d'après [21].

un examen par VCE de l'intestin grêle [20]. L'élaboration d'un tel score de propreté, idéalement automatisé, objectif, reproductible, avec de bonnes performances pour distinguer une préparation « adéquate » vs « inadéquate » est un enjeu de soins et de recherche.

/// **L'intelligence artificielle est au chevet du « score de Boston » pour le grêle** ///

En adaptant ce principe à l'échelle d'images fixes de VCE pour éviter des pertes d'information, notre équipe a développé un score colorimétrique électronique, basé sur le ratio de pixels rouges sur verts (R/V), dont les performances sont élevées par rapport à une lecture humaine experte : pour un ratio R/V = 1,6, la sensibilité et la spécificité sont supérieures à 90 % pour distinguer des images dont la qualité est « adéquate » ou « inadéquate » [21] (figure 6). Pour compléter cette approche,

notre équipe a développé un autre algorithme électronique automatisé afin de trier les images contenant des bulles en grande (> 10 % de la surface) ou faible (< 10 %) abondance (figure 7). En 37 millisecondes (ms) par image, la sensibilité et la spécificité de cet algorithme sont supérieures à 95 % en comparaison de lecteurs experts [22]. Un troisième paramètre (la luminosité de l'image) a été combiné au ratio R/V et à l'abondance des bulles, apportant de meilleures performances que chaque critère considéré isolément, avec une sensibilité de 90 % et une spécificité de 87 %, en 34 ms en moyenne [23]. Sur la base de ces travaux à l'échelle d'images fixes (brevet européen n°18305275.2-1124), nous approchons la question de l'évaluation algorithmique de la qualité de préparation des VCE, à l'échelle clinique, c'est-à-dire à l'échelle vidéo, à partir des données de l'essai randomisé contrôlé Prepintest [24]. Ce développement est un exemple désormais rare d'IA « artisanale » (*handcrafted*), et non basée sur l'apprentissage profond (*deep learning*).

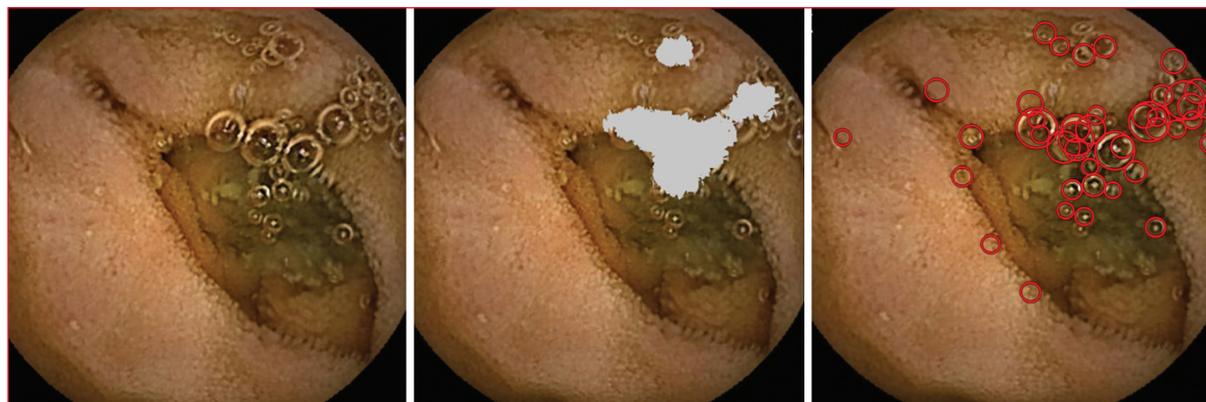
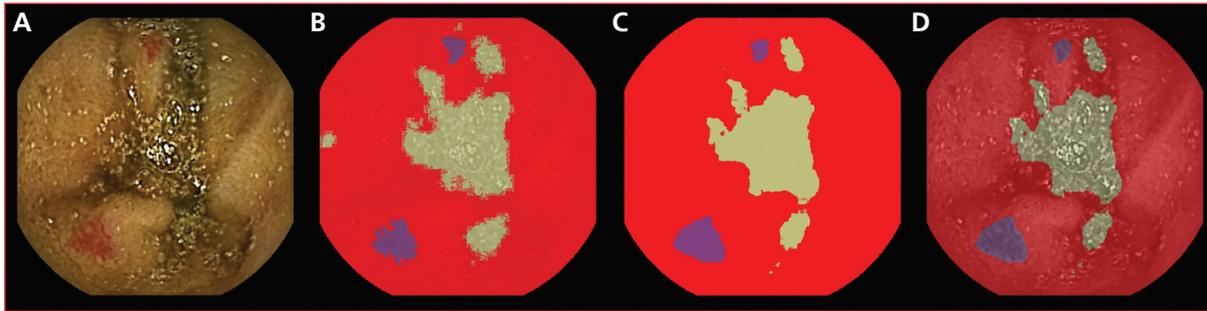


Figure 7 • Exemples de deux méthodes artisanales (*handcrafted methods*) pour la classification d'images de vidéocapsule endoscopique selon leur contenu en bulles, d'après [22].



**Figure 8** • Segmentation sémantique d'images de vidéocapsule par une approche de type deep learning. Cet exemple illustre la capacité d'une approche de type « réseaux neuronaux » à effectuer de manière précise une tâche de segmentation sémantique du contenu d'une image de vidéocapsule du grêle. Trois classes de pixels sont considérées : le tissu sain, le tissu pathologique (ici, une angiectasie), et des éléments perturbateurs (des bulles). L'image A correspond à l'image originale, l'image B à la « vérité terrain » déterminée par un clinicien, l'image C au résultat de segmentation automatique, et l'image D à la superposition image/segmentation automatique.

### • Aide à la détection de lésions en vidéocapsule endoscopique

L'IA peut aider à la détection d'anomalie lors de la lecture d'une VCE. La première étape de ce développement est la création de banques d'images, pour l'apprentissage machine. La banque CAD-CAP (*Computed-Assisted Diagnosis for Capsule Endoscopy*) est une banque multicentrique nationale d'images et de séquences vidéo courtes du grêle, anonymisées, normales et pathologiques, issues de VCE de troisième génération. Les lésions sélectionnées au sein d'images pathologiques sont délimitées et validées par un comité scientifique. L'ambition est de servir de « vérité terrain » pour le développement d'outils de DAO. À ce jour, 12 centres français ont fourni 4 796 examens du grêle, dont 3 712 normaux, et 1 084 pathologiques, parmi lesquelles 20 000 images normales et plus de 5 000 images de lésions étaient extraites, validées, classées par pertinence, et délimitées.

#### ■ Une Banque Nationale de milliers d'images de l'intestin grêle pour servir l'intelligence artificielle ■

À partir de la banque CAD-CAP, nous avons développé un algorithme diagnostique sur la base à la fois des méthodes « artisanales » (avec ACP) et des méthodes d'apprentissage profond (réseaux de neurones). L'algorithme était capable de distinguer des images normales d'images contenant une angiectasie typique (*figure 8*), avec une sensibilité de 100 % (aucune angiectasie n'était omise), une spécificité 95,83 %, en 47 ms par image en moyenne (et donc en extrapolant, moins de 40 minutes pour 50 000 images) [25]. La banque CAD-CAP a par ailleurs alimenté deux concours internationaux (l'un à Québec en 2017, l'autre à Grenade en 2018) où se sont affrontées plusieurs équipes de spécialistes en IA et en analyse d'images<sup>1</sup>. Lors de ces « challenges », nos résultats ont

été reproduits par plusieurs équipes, sur des lésions vasculaires plus subtiles encore (tâches érythémateuses, points rouges, angiectasies...) et sur des lésions inflammatoires ou ulcérées (érosions aphtoïdes, ulcérations superficielles ou profondes...).

#### ■ Les réseaux neuronaux interpréteront bientôt correctement les vidéocapsules en quelques minutes ■

### Vers une lecture automatisée de vidéocapsule endoscopique ?

Notre vision du développement du DAO à la lecture des VCE est que, dans un avenir proche, des centres équipés puissent transférer via des réseaux sécurisés, le soir même de leur enregistrement, les vidéos recueillies. Ces vidéos parviendront à des sites agréés, hébergeurs de données de santé, où des systèmes d'IA analyseront les images de façon exhaustive, relativement rapidement (non nécessairement en temps réel), avec une reproductibilité parfaite, et avec de bonnes performances diagnostiques. Ces systèmes pourront livrer une pré-lecture au médecin, notamment en termes de propreté (« validant » par exemple la qualité d'un examen normal), mais aussi en proposant des diagnostics (type, imputabilité et localisation des lésions), voire des options thérapeutiques (*figure 9*).

### Conclusion

L'intelligence artificielle a une place toute particulière dans l'analyse des signaux en général, et dans celle des images en particulier. De la même façon qu'elle se développe à grande vitesse dans tous les secteurs de notre vie, elle connaît une expansion importante dans le domaine des soins. Elle devrait donc impacter fortement les secteurs de l'imagerie médicale, en radiologie, mais aussi en anatomopathologie et certainement en endoscopie digestive. Des systèmes relativement avancés ont été validés en vidéocoloscopie, et d'autres sont

<sup>1</sup> <https://giana.grand-challenge.org>

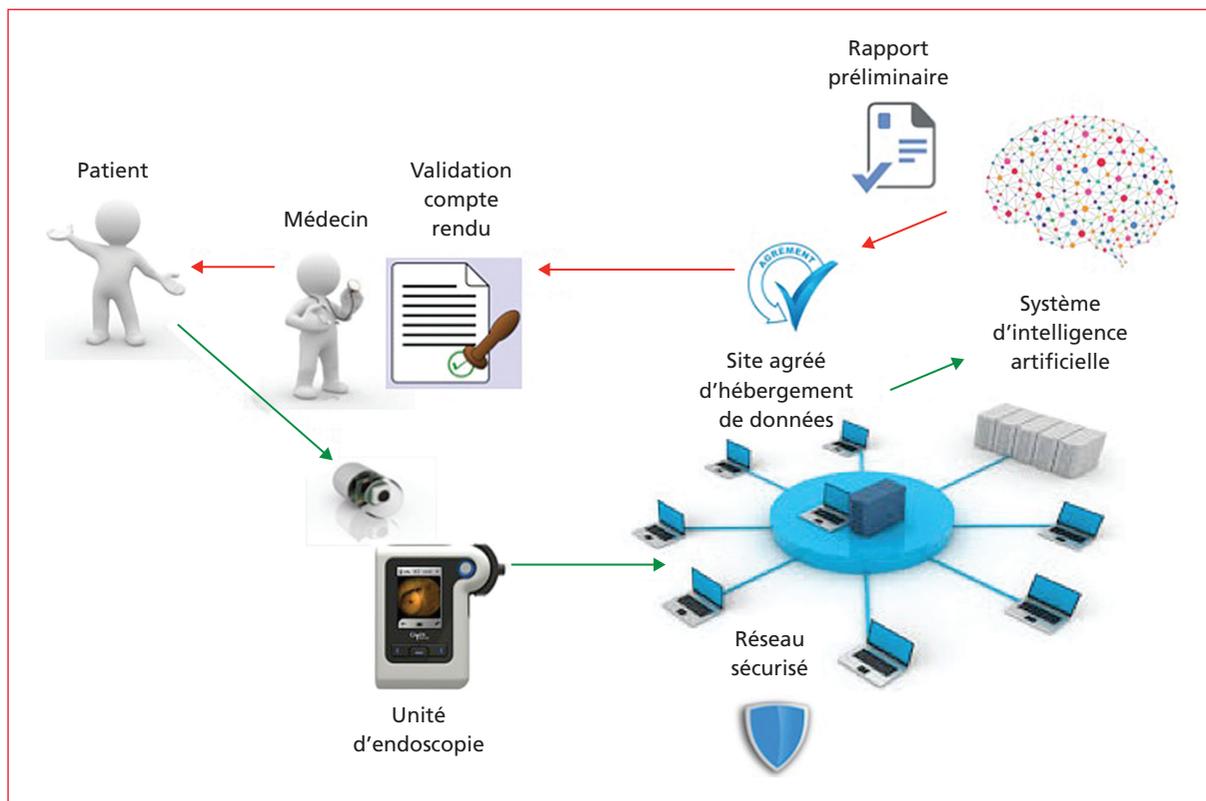


Figure 9 • Projet d'un système en réseau de lecture automatisée de vidéocapsules endoscopiques.

en développement en gastroscopie et en capsule endoscopique. La détection et la caractérisation de polypes coliques sont ses premières cibles, et les premières demandes de mise sur le marché sont en cours, mais des méthodes voient également le jour pour l'évaluation des maladies inflammatoires chroniques de l'intestin, de l'endobrachyoesophage, des tumeurs superficielles de l'estomac, des lésions kystiques du pancréas, etc. D'une technique à l'autre (endoscopie, capsule, échoendoscopie...) et d'une pathologie à l'autre, le phénomène et les utilisateurs devraient gagner en ampleur et confiance, respectivement. Aucune forme d'IA n'est pour l'instant industrialisée en endoscopie digestive, mais il fait peu de doute que tel sera le cas prochainement.

Cette révolution amène de nouvelles questions, non résolues pour l'instant. Quand et comment implémenter demain ces systèmes dans nos unités ? Avec quelle formation, à quel coût, et avec quel niveau de sécurité ? Ces systèmes renforceront-ils eux-mêmes l'apprentissage des machines ? Quelle sera alors la place du médecin en général, et de l'endoscopiste en particulier, dans cette médecine 2.0 ? L'IA restera-t-elle une simple « aide » ou s'autonomisera-t-elle dans certaines tâches ? Le praticien sera-t-il alors confiné à un rôle de simple exécutant... pour quelques années encore... ? ou devra-t-il valider les propositions des machines ?

Et surtout gagnera-t-il en noblesse en effectuant des tâches plus complexes ou plus « humaines » et éthiques que ne peuvent encore le faire des intelligences simulées ? Qu'en sera-t-il demain de la responsabilité d'un diagnostic ? Et si des machines en réseaux posent ces diagnostics, qui sera payé ? Les médecins ou les sociétés qui industrialiseront ces systèmes ? Quant aux gestes, il nous reste encore un peu de marge... les robots endoscopistes ne sont pas encore prêts, mais peut-être l'« ultra-intelligence » (l'IA créant elle-même des formes d'intelligence supérieure) les créera-t-elle...

**/// Les enjeux cliniques mais aussi éthiques, juridiques, et financiers sont majeurs ///**

D'ici là, si nous pensons que l'IA est utile à nos patients. Notre rôle dans son développement est majeur. Les réseaux neuronaux, par exemple, fers de lance de la plupart des systèmes d'IA en développement, existent depuis les années 1960. Il ne s'agit en rien d'une invention nouvelle. L'usage de l'IA devient courant parce que les puissances de calcul des processeurs augmentent considérablement, parce que les autoroutes de l'information permettent de leur transmettre des données, mais surtout parce que nous sommes en capacité de leur fournir les données qui les rendent « intelligents ». Sans bases de données, pas d'IA. Prioriser les projets, organiser

**TAKE HOME MESSAGES**

- L'intelligence artificielle est une science, et non plus une fiction.
- Ni médicament, ni dispositif, l'intelligence artificielle appliquée à la santé reçoit ses premières autorisations de mise sur le marché.
- Le traitement des images (endoscopiques notamment) se prête particulièrement bien aux applications d'intelligence artificielle. Tous les aspects (diagnostiques, thérapeutiques, pronostiques) de la discipline vont être impactés.
- La détection et la caractérisation de polypes coliques, assistés par ordinateur, devraient prochainement être commercialisées.
- L'intelligence artificielle représente une révolution technologique et soulève pour les sociétés savantes de nouvelles questions, éthiques et juridiques notamment.

les données, gérer leur qualité, c'est désormais une nouvelle tâche du médecin qui participe à l'effort de recherche clinique, via ces outils. Au fur et à mesure que les techniques d'imagerie endoscopique progressent (vidéocapsule, chromoendoscopies virtuelles, magnifications, microendoscopie confocale, endoscystoscopie, autofluorescence...), il faudra consacrer des « vérités » diagnostiques pour les machines, réalimenter leurs banques, évaluer cliniquement et économiquement l'apport de ces systèmes de DAO. Il reste donc beaucoup de travail pour nous, pauvres humains...

**Remerciements :**

Les auteurs remercient la SNFGE et MSD pour leur soutien dans le développement de la banque d'images CAD-CAP.

**Liens d'intérêts :**

X. Dray a été consultant pour Boston Scientific et Pentax, et est intervenu en formation pour Alfasigma, Bouchara-Recordati, Fujifilm, Medtronic et Norgine. R. Leenhardt, A. Histace et A. Becq déclarent n'avoir aucun lien d'intérêt en rapport avec l'article.

**Références**

Les références importantes apparaissent en gras.

- 1 • Li J, Rapin J, Rosier A, *et al.* Deep neural networks improve atrial fibrillation detection in Holter: first results. *Eur J Prevent Cardiol* 2016 ; 23(2S) : 41-55.
- 2 • **Gulshan V, Peng L, Coram M, *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016 ; 316 (22) : 2402-2410.**
- 3 • Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017 ; 542(7639) : 115-8.

- 4 • Haenssle HA, Fink C, Schneiderbauer R, *et al.* Man against machine : diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol* 2018 ; 29(8) : 1836-42.
- 5 • Litjens G, Kooi T, Bejnordi BE, *et al.* A survey on deep learning in medical image analysis. *Med Image Anal* 2017 ; 42 : 60-88.
- 6 • Stanek SR, Tavanapong W, Wong J, *et al.* SAPPHIRE: a toolkit for building efficient stream programs for medical video analysis. *Comput Methods Programs Biomed* 2013 ; 112 : 40421.
- 7 • Srinivasan N, Szewczynski M, Enders F, *et al.* Real-time feedback improves the quality of colonoscopy by trainees: a controlled clinical trial. *Am J Gastroenterol* 2012 ; 107 : S596.
- 8 • Silva JS, Histace A, Olivier Romain O, *et al.* Towards embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *Int J Comput Assist Radiol Surg* 2014 ; 9(2) : 283-93.
- 9 • Wang P, Xiao XX, Glissen Brown JR, *et al.* Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nature Biomed Eng* 2017 ; [Epub ahead of print].
- 10 • **Wang P, Bharadwaj S, Berzin TM, *et al.* Assistance of a real-time automatic colon polyp detection system increases polyp and adenoma detection during colonoscopy: a prospective randomized controlled study. United European Gastroenterology Week. 20-24 octobre 2018, Vienne, Autriche.**
- 11 • **Urbani G, Tripathi P, Alkayali T, *et al.* Deep Learning Localizes and Identifies Polyps in Real Time With 96 % Accuracy in Screening Colonoscopy. *Gastroenterology* 2018 ; 155 : 1069-1078.**
- 12 • Kominami Y, Yoshida S, Tanaka S, *et al.* Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016 ; 83 : 643-9.
- 13 • Misawa M, Kudo SE, Mori Y, *et al.* Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endoscopy. *Gastroenterology* 2016 ; 150 : 1531-2.
- 14 • Rath T, Tontini GE, Vieth M, *et al.* In vivo real-time assessment of colorectal polyp histology using an optical biopsy forceps system based on laser-induced fluorescence spectroscopy. *Endoscopy* 2016 ; 48 : 557-62.
- 15 • **Byrne MF, Chapados N, Soudan F, *et al.* Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2017.[Epub ahead of print].**
- 16 • Mori Y, Kudo S, Ogawa Y, *et al.* Can artificial intelligence correctly diagnose sessile serrated adenomas/polyps ? *Gastrointest Endosc* 2017 ; 85(5S) : AB510.
- 17 • Kanesaka T, Lee TC, Uedo N, *et al.* Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* 2018 ; 87(5) : 1339-44.
- 18 • Jacot GP, Qu C, Cahill J, *et al.* Use of artificial intelligence to predict future risk for improve CRC screening intervals. *Gastrointest Endosc* 2017 ; 85(5S) : AB734.
- 19 • Ichimasa K, Kudo SE, Mori Y, *et al.* Artificial intelligence may help in predicting the need for additional surgery after endoscopic resection of T1 colorectal cancer. *Endoscopy* 2018 ; 50 : 230-40.
- 20 • Brotz C, Nandi N, Conn M, *et al.* A validation study of 3 grading systems to evaluate small-bowel cleansing for wireless capsule endoscopy : a quantitative index, a qualitative evaluation, and an overall adequacy assessment. *Gastrointest Endosc* 2009 ; 69 : 262-70.
- 21 • Abou Ali E, Histace A, Camus M, *et al.* Development and validation of a computed assessment of cleansing score for evaluation of quality of small-bowel visualization in capsule endoscopy. *Endosc Int Open* 2018 ; 6 : E646-51.
- 22 • Pietri O, Rezgui G, Histace A, *et al.* Development and validation of an automated algorithm to evaluate the abundance of bubbles in small bowel capsule endoscopy. *Endosc Int Open* 2018 ; 6 : E462-9.
- 23 • Oumrani S, Histace A, Abou Ali E, *et al.* A multi-criterion, fully automated, high performance, rapid tool assessing the mucosal visualization quality of still images in small bowel capsule endoscopy. United European Gastroenterology Week 20-24 octobre 2018, Vienne, Autriche.
- 24 • Cholet F, Rahmi G, Gaudric M, *et al.* Does polyethylene glycol cleansing purge improve video capsule endoscopy diagnostic yield in obscure gastrointestinal bleeding. *Endoscopy* 2018 ; 50(4) : S18.
- 25 • **Leenhardt R, Vasseur P, Li C, *et al.* A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc.* 2018. [Epub ahead of print].**