

# La pathologie cancéreuse pulmonaire à l'heure de l'intelligence artificielle : entre espoir, désespoir et perspectives

Simon Heeke, Hervé Delingette, Youta Fanjat, Elodie Long-Mira, Sandra Lassalle, Véronique Hofman, Jonathan Benzaquen, Charles-Hugo Marquette, Paul Hofman, Marius Ilié

## ▶ To cite this version:

Simon Heeke, Hervé Delingette, Youta Fanjat, Elodie Long-Mira, Sandra Lassalle, et al.. La pathologie cancéreuse pulmonaire à l'heure de l'intelligence artificielle : entre espoir, désespoir et perspectives. Annales de Pathologie, 2019, 39 (2), pp.130-136. 10.1016/j.annpat.2019.01.003 . hal-02446712

## HAL Id: hal-02446712 https://inria.hal.science/hal-02446712

Submitted on 22 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



1	La Patnologie Cancereuse Pulmonaire à l'neure de l'Intelligence Artificielle :
2	Entre Espoir, Désespoir et Perspectives
3	The Age of Artificial Intelligence in Lung Cancer Pathology:
4	between Hope, Gloom and Perspectives
5	
6	Simon Heeke (1, 2), Hervé Delingette (3), Youta Fanjat (1), Elodie Long-Mira (1, 2), Sandra
7	Lassalle (1, 2), Véronique Hofman (1, 2), Jonathan Benzaquen (2, 4) Charles-Hugo Marquette
8	(4), Paul Hofman (1, 2), Marius Ilié (1, 2)
9	
10	(1) Laboratoire de Pathologie Clinique et Expérimentale/Biobanque (BB 0033-00025),
11	Fédération Hospitalo-Universitaire OncoAge, CHU de Nice, Université Côte d'Azur, 30 voie
12	Romaine, 06000 Nice, France
13	(2) Equipe 4, Institut de Recherche sur le Cancer et le Vieillissement de Nice (IRCAN),
14	Inserm U1081, CNRS UMR7284, Faculté de Médecine, 28, Avenue de Valombrose 06107
15	Nice, France
16	(3) Equipe Asclepios, INRIA Sophia-Antipolis, Université Côte d'Azur, 2004 Route des
17	Lucioles 06902 Sophia Antipolis, France
18	(4) Service de Pneumologie, Fédération Hospitalo-Universitaire OncoAge, CHU de Nice,
19	Université Côte d'Azur, 30 Voie Romaine, 06000 Nice, France
20	
21	Auteur correspondant : Marius Ilié, Laboratoire de Pathologie Clinique et Expérimentale,
22	Hôpital Pasteur, Pavillon J, 30 Voie Romaine 06000 Nice, France; Tél. 04 92 03 82 63; Fax.
23	04 92 03 85 55 ; ilie.m@chu-nice.fr
24	
25	
26	Mots clés: pathologie; histologie; cancer broncho-pulmonaire; intelligence artificielle;
27	réseaux de neurones convolutifs ; apprentissage profond
28	
29	Keywords: pathology; histology; lung cancer; artificial intelligence; convolutional neural
30	networks; deep learning
31	
32	
33	
34	IA et pathologie cancéreuse pulmonaire
	r

## Résumé

L'histopathologie est l'outil fondamental utilisé en anatomo-pathologie depuis plus d'un siècle pour établir le diagnostic final d'un carcinome broncho-pulmonaire. L'information phénotypique présente sur les images histologiques reflète l'effet global des altérations moléculaires sur le comportement des cellules cancéreuses et fournit une lecture visuelle pratique de l'agressivité de la maladie. Cependant, l'évaluation humaine de l'image histologique peut être parfois subjective et assez peu reproductible selon les cas. Par conséquent, l'analyse computationnelle de l'imagerie histologique via des approches dites « d'intelligence artificielle » (IA) a récemment reçu une attention considérable afin d'améliorer cette précision diagnostique. Ainsi, l'analyse computationnelle d'images de cancer du poumon a récemment été évaluée pour l'optimisation de la classification histologique ou cytologique, la prédiction du pronostic ou du profil génomique des patients atteints d'un cancer pulmonaire. Ce domaine, en pleine croissance, fait constamment preuve d'une grande puissance dans le domaine de l'informatique d'imagerie médicale en produisant des tâches de détection, de segmentation ou de reconnaissance d'une très grande précision. Cependant il subsiste plusieurs défis ou enjeux majeurs à relever afin de réussir le transfert réel de cette nouvelle approche en routine clinique. L'objectif de cette revue est de faire le point sur les applications récentes de l'IA en pathologie cancéreuse pulmonaire, mais aussi d'apporter des clarifications sur les avantages et les limites de cette approche, ainsi que les perspectives à mettre en œuvre pour un transfert potentiel dans la pratique quotidienne des pathologistes.

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

#### **Abstract**

Histopathology is the fundamental tool of pathology used for more than a century to establish the final diagnosis of lung cancer. In addition, the phenotypic data contained in the histological images reflects the overall effect of molecular alterations on the behavior of cancer cells and provides a practical visual reading of the aggressiveness of the disease. However, the human evaluation of the histological images is sometimes subjective and may lack reproducibility. Therefore, computational analysis of histological imaging using so-called "artificial intelligence" (AI) approaches has recently received considerable attention to improve this diagnostic accuracy. Thus, computational analysis of lung cancer images has recently been evaluated for the optimization of histological or cytological classification, prognostic prediction or genomic profile of patients with lung cancer. This rapidly growing field constantly demonstrates great power in the field of computing medical imaging by producing highly accurate detection, segmentation or recognition tasks. However, there are still several challenges or issues to be addressed in order to successfully succeed the actual transfer into clinical routine. The objective of this review is to emphasize recent applications of AI in pulmonary cancer pathology, but also to clarify the advantages and limitations of this approach, as well as the perspectives to be implemented for a potential transfer into clinical routine.

19 20

## Introduction

- 2 Environ 85% de tous les nouveaux cas de cancer du poumon sont classés comme des
- 3 carcinomes pulmonaires non à petites cellules (CNPC), 10% sont des carcinomes à petites
- 4 cellules (CPC) et 5% représentent les autres variantes [1]. Cette classification histo-
- 5 morphologique est un processus diagnostique clé car les options de traitement disponibles, y
- 6 compris les traitements conventionnels, la chimiothérapie et, plus récemment, les thérapies
- 7 ciblées et l'immunothérapie, différent considérablement en fonction du type histologique [1].
- 8 En effet, un diagnostic d'adénocarcinome pulmonaire entraînera la recherche de biomarqueurs
- 9 théranostiques et aura ainsi un impact important sur les options thérapeutiques [2].
- 10 L'interprétation précise de ces cancers peut être difficile et la distinction entre les différents
- 11 sous-types histologiques n'est pas toujours aisé, en particulier dans les tumeurs peu
- différenciées : dans ce cas, des études ancillaires sont indispensables pour une classification
- 13 plus précise [3].
- Depuis plus d'un siècle, l'histopathologie est l'outil clé de l'anatomo-pathologie, utilisée pour
- 15 établir le diagnostic final des carcinome broncho-pulmonaires. Les pathologistes évaluent des
- 16 caractéristiques histologiques telles que l'atypie nucléaire, l'activité mitotique, la densité
- cellulaire et l'architecture tissulaire, en intégrant les détails cytologiques, pour classifier et
- 18 grader les lésions cancéreuses. L'information phénotypique présente sur les images
- 19 histologiques reflète l'effet global des altérations moléculaires sur le comportement des
- cellules cancéreuses et fournit une lecture visuelle pratique de l'agressivité de la maladie [4].
- 21 Cependant, l'évaluation humaine de l'image histologique est parfois subjective et peu
- 22 reproductible. Par conséquent, l'analyse computationnelle de l'imagerie histologique via des
- 23 approches dites « d'intelligence artificielle » (IA) a récemment reçu une attention
- considérable pour améliorer cette précision diagnostique [5].
- De nombreux détails de la reconnaissance de l'image dans le système nerveux central sont
- 26 encore inconnus, mais nous pouvons déduire qu'au moins certaines, sinon la plupart, de nos
- 27 capacités cognitives utilisées grâce à notre cortex visuel et responsable du traitement des
- images, sont littéralement basées sur la « re connaissance ». Nous ré-identifions une image
- 29 que nous avons déjà vue et, selon la profondeur de la mémoire dans laquelle cette image est
- 30 stockée, nous la reconnaissons instantanément ou peu de temps après, avec un certain effort
- 31 mental (par exemple, lorsque nous sommes confrontés à un diagnostic que nous ne
- 32 connaissons pas bien ou que nous n'avons pas vu depuis plusieurs années). L'information sur
- les images, quel que soit le format sous lequel elle est stockée dans notre cerveau, fait

1 certainement l'objet de comparaisons et d'inférences sophistiquées à des fins d'identification 2 [6]. En analyse d'images médicales, il existe une large collection d'algorithmes informatiques qui 3 effectuent différentes opérations sur les images numériques : amélioration de la qualité, 4 filtrage, enregistrement et segmentation. Cette dernière opération a fait l'objet de recherches 5 approfondies pour quantifier la morphologie et la distribution des noyaux cellulaires [7]. 6 7 Aussi importantes que puissent être ces mesures, elles n'ont pas été en mesure d'entraîner un 8 changement significatif dans l'imagerie diagnostique, principalement parce que la 9 quantification conventionnelle est souvent introduite dans un algorithme "intelligent" pour produire une "classification" – une catégorie quelconque, généralement soit une décision 10 11 binaire (oui/non), soit une classification de la maladie. Ces quantifications n'ont pas 12 fondamentalement modifié le processus de diagnostic, probablement parce que de tels algorithmes informatiques ne réduisent pas l'incertitude pour accroître la confiance des 13 pathologistes dans un diagnostic précis [5]. Plus important encore, les algorithmes 14 informatiques axés sur la classification n'ont pas été en mesure d'aider vraiment les 15 pathologistes parce qu'ils ne fournissent aucun indice pour la rédaction du compte-rendu 16 17 histologique. Ainsi, la communauté des pathologistes s'est plutôt tournée vers le « deuxième avis », en particulier organisé par télépathologie pour réduire la variabilité entre les 18 19 observateurs [8]. La reconnaissance d'images, utilisée comme approche alternative à l'analyse d'images 20 médicales, offre la possibilité d'effectuer une "télépathologie virtuelle", en consultant d'autres 21 22 pathologistes et en accédant à leurs connaissances sans avoir besoin de leur présence 23 « physique » pour examiner les lames d'histologie. Elle nous permet également de consulter 24 non pas un seul pathologiste, mais autant que nous le souhaitons au sein d'un réseau d'experts de relecture (ex. MESOPATH) [9]. La reconnaissance d'images nous permet d'accéder à 25 26 l'expertise de plusieurs pathologistes en peu de temps et à un moindre coût qu'en le faisant par soi-même, ou même par la télépathologie en temps réel. Elle peut établir un cadre très fiable 27 28 pour évoluer vers le contrôle de la qualité grâce à l'établissement d'un consensus informatisé. 29 Mais pourquoi attribuons-nous de telles attentes à la reconnaissance d'images ? Bien que les 30 connexions synaptiques, avec leurs états binaires d'excitation et d'inhibition, soient les 31 éléments constitutifs du cerveau humain, l'inférence réelle est granulaire, floue, implicite et qualitative, par opposition à des caractéristiques spécifiques, certaines, explicites et 32

quantitatives, qui semblent nous permettre de traiter une information très complexe et

1 ambiguë comme les modèles tissulaires variables et les complexités du polymorphisme. Le 2 processus de diagnostic se termine par la rédaction d'un compte-rendu, une activité que l'on peut qualifier « d'informatique avec des mots ». La contradiction est que la communauté de la 3 vision par ordinateur et la communauté de l'IA entendent par "informatique" le simple fait de 4 comprimer et de produire des nombres. Nous pouvons ignorer le fonctionnement interne des 5 algorithmes d'IA, mais le résultat produit pourrait être décisif si cela améliore la performance 6 7 diagnostique des pathologistes ou renforce la confiance dans leurs conclusions. 8 Compte tenu du grand nombre de lésions diagnostiquées et des données générées (images et 9 données sur le traitement et le suivi), nous devrions être en mesure d'identifier et de récupérer des images qui sont histologiquement similaires à l'échantillon examiné, ainsi que les données 10 11 annotées pour chaque cas. Les comptes rendus contiennent les connaissances médicales de 12 nombreux autres pathologistes pour des cas similaires, ce qui en fait un trésor d'informations diagnostiques de haute qualité. Les logiciels informatiques de la prochaine génération 13 pourront rendre l'information brute directement accessible au pathologiste (en montrant les 14 15 images récupérées avec les rapports correspondants), ou ils peuvent fusionner l'information clé dans les rapports récupérés pour fournir le "sous-titrage automatique" de lames entières. 16 17 Ce dernier permettrait même de trier et de prioriser en temps réel les lames au fur et à mesure 18 qu'elles passent dans les scanners numériques. Le monde de la reconnaissance d'images basée 19 sur l'IA ouvre une vaste gamme d'options pour faire progresser et optimiser le flux de travail 20 du laboratoire de pathologie. En vision par ordinateur, l'extraction de caractéristiques visuelles consiste en des 21 22 transformations mathématiques calculées sur les pixels d'une image numérique. Les caractéristiques visuelles permettent généralement de mieux rendre compte de certaines 23 24 propriétés visuelles de l'image, utilisées pour des traitements ultérieurs entrant dans le cadre 25 d'applications telles que la détection d'objets ou la recherche d'images par le contenu. 26 La recherche d'images par le contenu (content-based image retrieval; CBIR) existe depuis plus de trois décennies [10]. Mais si nos attentes sont justifiées, pourquoi le CBIR n'a-t-il pas 27 28 tenu ses promesses ? Les raisons les plus importantes d'un point de vue technique, sont les 29 problèmes de calcul et d'exactitude. Dans le premier cas, il s'agit de la difficulté d'effectuer la 30 comparaison d'images en temps réel dans de grandes archives ; dans le second cas, il s'agit de 31 comparer correctement les images afin que les images identifiées soient réellement similaires 32 à l'image de la requête. Du point de vue de la pathologie digitale, les obstacles sont 33 légèrement différents. Ainsi, la principale raison pour laquelle les systèmes CBIR n'ont pas

1 réussi le transfert en routine clinique est liée très probablement à ce qu'on appelle le "fossé sémantique" [11]. Les représentations d'images en vision par ordinateur sont numériques et 2 objectives, alors que les pathologistes utilisent des représentations verbales et subjectives qui 3 souvent ne peuvent pas être modélisées ou analysées. L'écart qui en résulte entre les 4 ordinateurs et les experts pathologistes ne permet pas une définition sans ambiguïté de la 5 similarité. En fait, le fossé sémantique est sans doute le principal défi que pose l'adoption du 6 7 CBIR dans le déroulement du travail dans un laboratoire d'anatomo-pathologie; les résultats du CBIR n'ont pas été jusqu'à présent acceptables pour les pathologistes. Le chemin vers les 8 9 images récupérées n'est pas pertinent si le pathologiste n'est pas d'accord sur le fait que les images appariées sont vraiment similaires à l'image de la requête - une mauvaise réponse est 10 11 fausse, peu importe comment elle a été atteinte. Ces dernières années, cela a commencé à 12 changer et le CBIR connaît une renaissance avec la promesse de la révolution de l'IA [11]. L'IA est un terme général pour designer « l'ensemble des théories et des techniques mises en 13 œuvre en vue de réaliser des machines capables de simuler l'intelligence ». Souvent classée 14 15 dans le groupe des sciences cognitives, elle fait appel à la neurobiologie computationnelle, à 16 la logique mathématique et à l'informatique. 17 Depuis sa naissance, il y a 70 ans, avec quelques abstractions simples du fonctionnement d'un neurone dans le cerveau humain, l'IA est devenue un outil indispensable pour les applications 18 19 de vision par ordinateur. Plus particulièrement, les réseaux de neurones artificiels (artificial neural networks; ANN) ont gagné une grande popularité en raison de leur capacité de 20 21 reconnaissance impressionnante lorsqu'ils sont mis en œuvre avec de nombreuses couches de 22 neurones artificiels (unités de traitement qui peuvent effectuer une simple agrégation des 23 valeurs synaptiques entrantes provenant d'autres unités). Ces ANN "profonds" reconnaissent 24 le contenu d'une image numérique en apprenant une représentation compacte de l'image, « un codage élégant », que nous pouvons supposer être un modèle informatique primaire, mais 25 26 fonctionnel, au même titre qu'une image rétinienne quand elle traverse le nerf optique pour 27 atteindre le cortex visuel du cerveau humain [12]. 28 Les réseaux de neurones convolutifs (Convolutional Neural Networks; CNN) sont parmi les 29 solutions les plus efficaces pour extraire des caractéristiques pertinentes des images 30 numériques (Fig. 1). Un exemple typique est d'apprendre 1024 traits profonds pour 31 représenter une structure histologique représentée dans une image 240x240, réduisant l'information à moins de 2% de sa taille originale (Fig. 1). Pour créer de telles représentations 32 33 compactes, les réseaux profonds ajustent généralement plusieurs centaines de milliers de

- synapses artificielles pour atteindre leur objectif d'apprentissage, un processus de formation
- 2 dominé par l'essai et l'erreur dans la phase de conception et de nombreuses heures, voire des
- 3 jours de formation réelle. Un nombre croissant d'articles fait état d'une grande précision pour
- 4 la reconnaissance d'images histologiques numériques à l'aide de réseaux profonds [13-17].

6

## Applications de l'IA en pathologie cancéreuse pulmonaire

- 7 Pour aider les pathologistes, l'analyse automatique d'images des cancers du poumon a
- 8 récemment été évaluée pour l'optimisation de la classification histologique ou cytologique, la
- 9 prédiction du pronostic ou du profil génomique des patients [15, 18, 19].
- Yu et coll. ont combiné des techniques conventionnelles de seuillage et de traitement d'images
- avec des méthodes computationnelles d'apprentissage (« machine learning ») atteignant une
- 12 aire sous la courbe (AUC) ~0,85 pour distinguer le tissu normal du tissu néoplasique, et ~0,75
- dans la distinction entre les carcinomes épidermoïdes et les adénocarcinomes pulmonaires
- 14 [19]. Les auteurs ont également démontré que les caractéristiques morphologiques
- quantitatives extraites des lames histologiques peuvent prédire la survie des patients atteints
- de CNPC [19]. Plus récemment, l'apprentissage approfondi (« deep learning ») a été évalué
- pour la classification de tumeurs du sein, de la vessie et du poumon, atteignant une AUC de
- 18 0,83 pour la classification des sous-types histologiques de tumeurs pulmonaires sur des lames
- 19 virtuelles disponibles dans l'Atlas du Génome du Cancer (*The Cancer Genome Atlas, TCGA*)
- 20 [13]. De manière intéressante, la précision de prédiction des sous-types histologiques de
- 21 CNPC atteint ~0,94 grâce à l'analyse du niveau d'ADN plasmatique [20], alors que l'ajout de
- 22 marqueurs immunochimiques permet d'obtenir une AUC de ~0,941 [21]. Plus récemment,
- Coudray et coll. ont démontré que l'analyse CNN à partir de lames virtuelles intégrées dans
- TCGA serait capable de classer presque sans ambiguïté les tissus normaux par rapport aux
- 25 tissus tumoraux (AUC ~0,99) et pourrait distinguer les sous-types de CNPC avec une grande
- précision (AUC ~0,97), atteignant une sensibilité et une spécificité comparable à celle d'un
- 27 pathologiste [15]. Environ la moitié des cas classés par erreur par l'algorithme n'ont pas été
- 28 correctement classés par les pathologistes, mettant en évidence la difficulté intrinsèque de
- 29 distinguer un carcinome épidermoïde d'un adénocarcinome pulmonaire dans certains cas. De
- 30 manière intéressante, 83% de cas mal classés par au moins un des pathologistes a été
- 31 correctement assigné par l'algorithme, ce qui suggère que ce modèle pourrait accompagner les
- pathologistes dans leur diagnostic [15]. Cependant, les images virtuelles utilisées dans cette
- 33 étude pour entraîner l'algorithme ne sont pas représentatives de la diversité et de

- 1 l'hétérogénéité des tissus (e par exemple les foyers de nécrose, la présence des vaisseaux
- 2 sanguins et/ou d'une inflammation) que les pathologistes analysent habituellement. Enfin,
- 3 cette étude met les bases de la prédiction du statut mutationnel des gènes EGFR, STK11,
- 4 FAT1, SETBP1, KRAS et TP53 à partir l'analyse CNN d'images virtuelles, avec notamment
- 5 une grande précision pour la prédiction de la présence des mutations de STK11 (AUC ~0,85)
- 6 [15]. Néanmoins, l'algorithme doit être encore amélioré pour la prédiction du statut
- 7 mutationnel de l'EGFR (AUC ~0,75) [15].
- 8 Un algorithme CNN a été récemment évalué comme outil d'aide diagnostique pour la
- 9 classification des sous-types de CNPC à partir d'images virtuelles de cytologie, démontrant
- une précision diagnostique qui cependant nécessite d'être optimisée (AUC ~0,70) [22].

12

#### Limites actuelles de l'IA

- 13 L'IA est un domaine en pleine croissance, en train de devenir un outil de pointe pour la vision
- par ordinateur et la reconnaissance automatisée d'images. Il fait constamment preuve d'une
- 15 grande puissance dans le domaine de l'informatique d'imagerie médicale en produisant des
- tâches de détection, de segmentation ou de reconnaissance d'une grande précision [5]. Malgré
- 17 le succès rencontré dans la littérature, il subsiste à ce jour plusieurs défis importants ou enjeux
- 18 à relever afin de réussir le transfert réel en routine clinique.

- 20 Limites technologiques
- 21 La façon dont les algorithmes obtiennent d'excellentes performances n'est pas complètement
- et théoriquement comprise [23].La grande quantité de données d'images microscopiques
- 23 nécessite un taux élevé de traitement informatique, nécessitant ainsi une accélération des
- calculs en CNN, mais aussi une grande capacité de stockage d'images à haute définition. Les
- 25 CNN n'ont pas encore actuellement saisi complétement les aspects visuels fournis par les
- 26 images, et cela affecterait la capacité d'interprétation des données d'imagerie qui fournissent
- une très grande richesse pour la caractérisation d'une pathologie [5].
- Un autre défi dans l'analyse d'images microscopiques est celui lié aux artefacts d'images et
- 29 des effets de lots. L'acquisition d'images microscopiques est très différente de celle utilisée
- dans l'imagerie médicale, où l'anatomie est à peu près dans la même position à une échelle
- 31 identique. En histologie, les images microscopiques sont acquises après une série d'étapes de
- 32 traitement, comme le prélèvement, la fixation, l'inclusion en paraffine, la coupe et la
- coloration [24]. Des erreurs à tous ces niveaux peuvent entraîner des artefacts sur l'image

1 finale tels que des plis, des ombres, des zones floues et une variation de coloration [25]. Ces 2 résultats indésirables affectent considérablement la qualité de l'image et remettent en question son calcul automatique, y compris l'apprentissage en profondeur. Bien qu'il existe quelques 3 études sur la correction d'artefacts d'images microscopiques, la littérature sur ce sujet est 4 rare[24, 25]. Les effets d'échelle par lots sont plus difficiles à détecter et à traiter, car l'échelle 5 de l'objet peut naturellement changer en raison de la progression de la maladie, ce qui ne 6 7 facilite pas la distinction avec celle des procédures d'imagerie très disparates. Ces effets de lot peuvent générer un biais dans l'évaluation de la performance des modèles prédictifs [5]. Par 8 conséquent, les algorithmes développés pour le calcul d'images histologiques doivent être 9 robustes par rapport à ces variations au sein d'un même centre mais aussi entre les différents 10 11 centres. L'IA progresse à un tel rythme qu'elle dépasse de loin ce qu'un seul fournisseur de technologie 12 ne pourrait jamais espérer couvrir. Les fournisseurs de technologies d'imagerie doivent fournir 13 une interface ouverte pour offrir aux pathologistes la possibilité d'appliquer leurs innovations. 14 15 Les pathologistes ont l'habitude de travailler avec une seule interface utilisateur harmonisée, 16 qui intègre toutes les informations diagnostiques à portée de main. Si elle devait être utilisée 17 et acceptée par la communauté hospitalière, l'IA doit être mise à la disposition des utilisateurs 18 dans les systèmes déjà utilisés, tels que leur système d'archivage et de transmission d'images, 19 afin qu'ils puissent accéder aisément et rapidement à la fonctionnalité d'IA pour l'imagerie.

20

22

23

24

25

26

27

28

29

30

31

32

33

21 Responsabilité(s) en e-Pathologie

L'avènement de l'IA en pathologie marque de nouveaux espoirs mais aussi de nouvelles contraintes pour les anatomo-pathologistes. D'un côté, nous pouvons nous attendre à profiter de divers progrès avec par exemple comme conséquence, une diminution de l'erreur diagnostique et implicitement l'augmentation du taux de reproductibilité inter-pathologistes, et donc une diminution à *priori* des sources de responsabilité. Mais d'un autre côté, il existe une contrepartie aux progrès attendus dans ce domaine. En effet si le résultat attendu n'est pas atteint par la technologie, il se peut qu'il y ait une augmentation du nombre de recours en responsabilité [26]. Nous sommes naturellement plus sévères vis-à-vis du résultat obtenu à l'aide d'une machine que de celui obtenu par l'homme à qui l'on confère « un droit à l'erreur » [27]. L'introduction croissante de cette technologie nécessite d'appréhender les risques de défaillance et donc une problématique de responsabilité : en cas d'incident, qui est responsable [26, 27] ?

- 2 Enseigner l'IA aux futures générations de pathologistes
- 3 A court terme, il est envisageable de réutiliser des réseaux CNN pré-entrainés sur d'autres
- 4 problématiques [15], par des techniques de « transfer learning ». Des librairies d'images
- 5 « open-source » permettent de coder cela en quelques lignes. A l'avenir, afin de déjouer les
- 6 embuches théoriques et pratiques, des bases simples de codage devront être enseignés aux
- 7 jeunes pathologistes, pouvant ouvrir la voie aux « pathologistes computationnels ».

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

## Perspectives

La grande majorité d'articles scientifiques traitant de l'IA en pathologie cancéreuse pulmonaire utilise des traits profonds à des fins de classification (par exemple, si une image représente ou non une tumeur maligne). Les solutions de recherche d'images basées sur le CBIR médical s'abstiennent de cette approche. Le CBIR médical est fondamentalement centré sur le pathologiste, contrairement à l'IA fondée sur la classification, qui vise essentiellement à prendre des décisions au nom du pathologiste. Il est compréhensible que les pathologistes soient initialement opposés à cette dernière solution, mais la première fait un usage précieux des solutions d'IA. Au lieu de laisser les CNN et autres ANN profondes utiliser les représentations d'images extraites (caractéristiques profondes) comme base pour une classification binaire ("oui/non") du cancer, nous pouvons les utiliser pour indexer et récupérer des images de lames virtuelles entières, ce qui engendre plusieurs avantages. Premièrement, les capacités de reconnaissance d'images des réseaux profonds ont montré empiriquement que l'écart sémantique entre les perceptions informatiques et humaines peut être comblé [11]. Deuxièmement, l'IA offre une multitude de techniques polyvalentes de reconnaissance, d'indexation et de recherche. Enfin les progrès des logiciels et du matériel ont permis d'effectuer des millions de comparaisons d'images en une fraction de seconde [5]. Le fait que nous vivions actuellement une transition de la microscopie à la pathologie digitale n'est qu'une coïncidence étonnante qui profite à l'adoption de la vision par ordinateur en pathologie. L'un des domaines qui profitera le plus du transfert de l'IA en clinique est certainement l'immuno-oncologie. Ainsi, l'évaluation computationnelle de l'expression de PD-L1 pourrait aider le pathologiste à surmonter les obstacles liés à un marquage hétérogène pouvant conduire une mauvaise stratification pour le traitement, en fournissant une mesure normalisée à la résolution d'une seule cellule sur coupe tissulaire entière [28]. L'IA pourrait améliorer l'évaluation de l'infiltrat immunitaire intra-tumoral (« Cancer Immunogram ») pour

1 permettre une meilleure stratification des patients inclus dans des essais cliniques en immuno-2 oncologie [29]. Pour permettre une stratification plus précise des patients devant bénéficier de l'immunothérapie, l'analyse histologique devrait viser une caractérisation simultanée des voies 3 immunitaires et tumorales dans un seul échantillon tissulaire [30-32]. L'immunoprofilage 4 multiplexé est essentiel pour générer des données complètes sur des biomarqueurs potentiels 5 afin d'assurer une corrélation plus précise avec les paramètres cliniques [28]. Dans ce 6 7 contexte, l'IA en pathologie digitale est un outil efficace pour extraire des informations complètes sur les niveaux d'expression des biomarqueurs, leur co-localisation et leur 8 9 représentation spatiale. Lorsque l'on s'intéresse aux chiffres clés de l'anatomie pathologique, de l'effectif national en 10 11 passant par la démographie par région ou encore l'âge moyen par sexe, l'intégration de l'IA pourrait permettre l'amélioration de la capacité du système national de santé à orchestrer son 12 flux de travail et à permettre un triage plus automatique des images histologiques et 13 cytologiques, avec une automatisation des lectures répétitives, afin que les cas difficiles et 14 urgents puissent être traités en priorité. 15 16 L'idée que 100% de l'histologie soit analysée et rapportée par des machines d'IA, plutôt que 17 par une personne, est très éloignée de la réalité et pourra certainement ne jamais se produire, 18 mais cela ne veut pas dire qu'il n'y a pas certains bénéfices à en tirer très rapidement. Avec un 19 besoin urgent de répondre à des volumes de travail croissants et le fait qu'il y ait tout simplement de moins en moins de professionnels pour faire face à une demande croissante, 20 l'IA offre la possibilité d'éviter des retards coûteux dans les diagnostics qui pourraient nuire 21 22 considérablement à la prise en charge des patients. 23 En dépit des opportunités évidentes, il existe encore de nombreux obstacles à surmonter si 24 nous voulons amener les systèmes CBIR dans les laboratoires de pathologie - en particulier la nécessité d'une validation approfondie et complète de la recherche d'images à différentes fins 25 26 en pathologie. Contrairement à la classification des images, qui peut être validée en 27 laboratoire d'ingénierie, la recherche d'images ne peut pas être validée sans la présence et 28 l'implication très active des pathologistes. Il y a une lueur d'espoir car la technologie met 29 l'accent sur les pathologistes, plutôt que de chercher à les remplacer. Les systèmes CBIR 30 existent pour aider les pathologistes - et ils ne peuvent être conçus et validés sans notre

participation directe. De plus, une fois utilisés, ils ne peuvent pas continuer à « apprendre »

sans que les pathologistes soient au cœur de ce processus. L'IA ne remplacera pas les

31

- pathologistes; pourtant, les pathologistes qui utilisent l'IA pourraient remplacer ceux qui ne
- 2 l'utilisent pas.
- 3 La conception, la validation et l'approbation réglementaire des solutions de recherche
- 4 d'images ne se feront certainement pas du jour au lendemain. En attendant, nous pouvons
- 5 identifier des cas d'utilisation pratique pour la recherche d'images qui démontrent comment
- 6 elle peut nous propulser vers l'établissement d'un consensus informatique. Avec le succès
- 7 récent de l'IA dans une multitude d'applications de vision par ordinateur et la croissance
- 8 rapide de la pathologie digitale, nous nous rapprochons de plus en plus de l'horizon des
- 9 partenariats pathologistes-ordinateurs.
- 10 En conclusion, l'utilisation de l'IA progresse de plus en plus dans le domaine de la recherche
- translationnelle en pathologie cancéreuse pulmonaire, mais n'a pas encore été mise en œuvre
- de la manière la plus pratique pour le transfert en routine clinique. Une application utile de
- 13 l'IA qui pourrait être rapidement déployée consiste en la recherche et la récupération d'images,
- 14 une tâche que les ordinateurs peuvent effectuer beaucoup plus rapidement que les
- pathologistes. De nouvelles approches utilisant des réseaux neuronaux artificiels peuvent
- aider à surmonter les défis de la reconnaissance d'images par ordinateur. L'extraction d'images
- basée sur le contenu peut s'appuyer sur l'IA, tout en restant centrée sur le pathologiste, car ne
- pouvant fonctionner sans la contribution d'une participation humaine.

## Références

- 2 1. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization
- 3 Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since
- 4 the 2004 Classification. J Thorac Oncol. 2015;10:1243-60.
- 5 2. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. Science.
- 6 2018;359:1350-5.
- 7 3. Zachara-Szczakowski S, Verdun T, Churg A. Accuracy of classifying poorly
- 8 differentiated non-small cell lung carcinoma biopsies with commonly used lung carcinoma
- 9 markers. Hum Pathol. 2015;46:776-82.
- 10 4. Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from
- 11 histology and genomics using convolutional networks. Proc Natl Acad Sci U S A.
- 12 2018;115:E2970-E9.
- 5. Fuyong X, Yuanpu X, Hai S, et al. Deep Learning in Microscopy Image Analysis: A
- 14 Survey. IEEE Trans Neural Netw Learn Syst. 2018;29:4550-68.
- 15 6. Jia K, Xue X, Lee JH, et al. Visual perceptual learning modulates decision network in
- the human brain: The evidence from psychophysics, modeling, and functional magnetic
- 17 resonance imaging. J Vis. 2018;18:9.
- 18 7. Sirinukunwattana K, Ahmed Raza SE, Yee-Wah T, et al. Locality Sensitive Deep
- 19 Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology
- 20 Images. IEEE Trans Med Imaging. 2016;35:1196-206.
- 8. Farahani N, Pantanowitz L. Overview of Telepathology. Clin Lab Med. 2016;36:101-
- 22 12.
- 9. Galateau Salle F, Le Stang N, Nicholson AG, et al. New Insights on Diagnostic
- 24 Reproducibility of Biphasic Mesotheliomas: A Multi-Institutional Evaluation by the
- 25 International Mesothelioma Panel From the MESOPATH Reference Center. J Thorac Oncol.
- 26 2018;13:1189-203.
- 27 10. Sparks R, Madabhushi A. Out-of-Sample Extrapolation utilizing Semi-Supervised
- 28 Manifold Learning (OSE-SSL): Content Based Image Retrieval for Histopathology Images.
- 29 Sci Rep. 2016;6:27306.
- 30 11. Traore L, Kergosien Y, Racoceanu D. Bridging the Semantic Gap Between Diagnostic
- 31 Histopathology and Image Analysis. Stud Health Technol Inform. 2017;235:436-40.
- 32 12. Liu Y, Kohlberger T, Norouzi M, et al. Artificial Intelligence-Based Breast Cancer
- Nodal Metastasis Detection. Arch Pathol Lab Med. 2018.

- 1 13. Khosravi P, Kazemi E, Imielinski M, et al. Deep Convolutional Neural Networks
- 2 Enable Discrimination of Heterogeneous Digital Pathology Images. EBioMedicine.
- 3 2018;27:317-28.
- 4 14. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic Assessment of
- 5 Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast
- 6 Cancer. JAMA. 2017;318:2199-210.
- 7 15. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation
- 8 prediction from non-small cell lung cancer histopathology images using deep learning. Nat
- 9 Med. 2018;24:1559-67.
- 10 16. Mezheyeuski A, Bergsland CH, Backman M, et al. Multispectral imaging for
- 11 quantitative and compartment-specific immune infiltrates reveals distinct immune profiles
- that classify lung cancer patients. J Pathol. 2018;244:421-31.
- 13 17. Qiu JX, Yoon HJ, Fearn PA, et al. Deep Learning for Automated Extraction of
- Primary Sites From Cancer Pathology Reports. IEEE J Biomed Health Inform. 2018;22:244-
- 15 51.
- 16 18. Luo X, Zang X, Yang L, et al. Comprehensive Computational Pathological Image
- Analysis Predicts Lung Cancer Prognosis. J Thorac Oncol. 2017;12:501-9.
- 19. Yu KH, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by
- 19 fully automated microscopic pathology image features. Nat Commun. 2016;7:12474.
- 20 20. Sozzi G, Conte D, Leon M, et al. Quantification of free circulating DNA as a
- 21 diagnostic marker in lung cancer. J Clin Oncol. 2003;21:3902-8.
- 22 21. Terry J, Leung S, Laskin J, et al. Optimal immunohistochemical markers for
- 23 distinguishing lung adenocarcinomas from squamous cell carcinomas in small tumor samples.
- 24 Am J Surg Pathol. 2010;34:1805-11.
- 25 22. Teramoto A, Tsukamoto T, Kiriyama Y, et al. Automated Classification of Lung
- 26 Cancer Types from Cytological Images Using Deep Convolutional Neural Networks. Biomed
- 27 Res Int. 2017;2017:4067832.
- 28 23. Nitta T. Resolution of Singularities Introduced by Hierarchical Structure in Deep
- Neural Networks. IEEE Trans Neural Netw Learn Syst. 2017;28:2282-93.
- 30 24. Farris AB, Cohen C, Rogers TE, et al. Whole Slide Imaging for Analytical Anatomic
- 31 Pathology and Telepathology: Practical Applications Today, Promises, and Perils. Arch
- 32 Pathol Lab Med. 2017;141:542-50.

- 1 25. Kothari S, Phan JH, Stokes TH, et al. Pathology imaging informatics for quantitative
- analysis of whole-slide images. J Am Med Inform Assoc. 2013;20:1099-108.
- 3 26. Winfield AFT, Jirotka M. Ethical governance is essential to building trust in robotics
- 4 and artificial intelligence systems. Philos Trans A Math Phys Eng Sci. 2018;376.
- 5 27. Courtland R. Bias detectives: the researchers striving to make algorithms fair. Nature.
- 6 2018;558:357-60.
- 7 28. Koelzer VH, Sirinukunwattana K, Rittscher J, et al. Precision immunoprofiling by
- 8 image analysis and artificial intelligence. Virchows Arch. 2018.
- 9 29. Blank CU, Haanen JB, Ribas A, et al. CANCER IMMUNOLOGY. The "cancer
- 10 immunogram". Science. 2016;352:658-60.
- 11 30. Gerdes MJ, Sevinsky CJ, Sood A, et al. Highly multiplexed single-cell analysis of
- formalin-fixed, paraffin-embedded cancer tissue. Proc Natl Acad Sci U S A. 2013;110:11982-
- 13 7.
- 14 31. Ilie M, Beaulande M, Ben Hadj S, et al. Chromogenic Multiplex
- 15 Immunohistochemistry Reveals Modulation of the Immune Microenvironment Associated
- with Survival in Elderly Patients with Lung Adenocarcinoma. Cancers (Basel). 2018;10.
- 17 32. Ilie M, Beaulande M, Hamila M, et al. Automated chromogenic multiplexed
- immunohistochemistry assay for diagnosis and predictive biomarker testing in non-small cell
- 19 lung cancer. Lung Cancer. 2018;124:90-4.

## Déclaration d'intérêts

2 Les auteurs déclarent ne pas avoir de conflits d'intérêts en relation avec cet article.

3 4

1

5

6 7

11

12

13

14

15

16

17

Figure 1. Algorithme diagnostique schématisé en pathologie cancéreuse broncho-8 9 pulmonaire. A) L'image numérique d'une lame HES montrant les caractéristiques 10

histologiques d'un adénocarcinome pulmonaire peu différencié. B) Les algorithmes de

classification basés sur l'analyse histologique et moléculaire des pathologistes utilisent des approches simples et hautement reproductibles basées sur des arbres de décision. C) Les

approches fondées sur un algorithme IA/CNN peuvent permettre l'extraction et l'association

de caractéristiques multiparamétriques qui facilitent même l'intégration de groupes subtils

mais reproductibles de caractéristiques à des systèmes de classification. Les CNN peuvent

donc tolérer l'extraction fiable d'un plus grand nombre de caractéristiques non prises en

compte par l'œil humain qui peuvent être corrélées à des entités, des résultats et des

altérations moléculaires spécifiques. 18

