



HAL
open science

Intelligence artificielle et radiothérapie : quelles bases et quelles perspectives ?

A. Burgun

► **To cite this version:**

A. Burgun. Intelligence artificielle et radiothérapie : quelles bases et quelles perspectives?. *Cancer/Radiothérapie*, 2019, 23, pp.913 - 916. 10.1016/j.canrad.2019.08.005 . hal-03488991

HAL Id: hal-03488991

<https://hal.science/hal-03488991>

Submitted on 21 Jul 2022

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Intelligence artificielle et radiothérapie : quelles bases et quelles perspectives ? *

Basis and perspectives of artificial intelligence in radiation therapy

Anita BURGUN ^{a,b*}

^a Équipe Sciences de l'information pour la médecine personnalisée, centre de recherche des Cordeliers, université Paris Descartes, université de Paris, 15 rue de l'École-de-Médecine, 75006 Paris

^b Service d'informatique médicale, AP-HP hôpital européen Georges-Pompidou, 15, rue Leblanc, 75015 Paris

E-mail : anita.burgun@aphp.fr

*Cet article a fait l'objet d'une conférence lors du congrès de la Société française de radiothérapie oncologique le vendredi 11 octobre 2019.

Résumé

L'intelligence artificielle est une notion hautement polysémique. Pour réaliser un raisonnement complexe dans la vie réelle, et s'adapter à des connaissances et des situations nouvelles, deux grandes approches sont développées en informatique: les réseaux de neurones basés sur le modèle connexionniste (*deep learning*) pour l'apprentissage, et les méthodes symboliques et logiques capables de travailler à un niveau abstrait de description et de raisonnement. Les algorithmes d'intelligence artificielle reproduisant les processus de déduction, induction et abduction ont des applications en radiothérapie. Combinés à la radiomique, les réseaux de neurones ont obtenu de bons résultats en classification d'images, traitement du langage naturel, phénotypage à partir des dossiers patients, adaptation des traitements. Les approches logiques ont produit des ontologies formelles, des algorithmes déterministes pour la décision et des méthodes de vérification de cohérence des systèmes complexes. Une intelligence artificielle hybride conjuguant apprentissage et logique est nécessaire pour réaliser des tâches complexes allant au delà de l'intelligence artificielle qui réalise des tâches restreintes et spécialisées. Combinée à des modèles formalisant les connaissances physico-biologiques, l'intelligence artificielle est au cœur de nouveaux outils comme les jumeaux numériques (*digital twins*) nécessaires à la médecine de précision en oncologie.

Abstract

Artificial intelligence is a highly polysemic term. In computer science, with the objective of being able to solve totally new problems in new contexts, artificial intelligence includes connectionism (neural networks) for learning and logics for reasoning. Artificial intelligence algorithms mimic tasks normally requiring human intelligence, like deduction, induction, and abduction. All apply to radiation oncology. Combined with radiomics, neural networks have obtained good results in image classification, natural language processing, phenotyping based on electronic health records, and adaptive radiation therapy. General adversarial networks have been tested to generate synthetic data. Logics based system have been developed for providing formal domain ontologies, supporting clinical decision and checking consistency of the systems. Artificial intelligence must integrate both deep learning and logic approaches to perform complex tasks and go beyond the so-called narrow artificial intelligence that is tailored to perform some highly specialized task. Combined together with mechanistic models, artificial intelligence has the potential to provide new tools such as digital twins for precision oncology.

Mots clé

intelligence artificielle ; oncologie ; radiothérapie ; médecine de précision ; algorithme

Keywords

artificial intelligence; precision oncology; radiation therapy; algorithm

1. Introduction

L'intelligence artificielle est une notion hautement polysémique. Elle doit inclure plusieurs processus « intelligents » pour réaliser un raisonnement complexe et adapté face à une question de la vie réelle, et s'adapter à des connaissances et des situations nouvelles. Pour réussir cela, deux grandes méthodes complémentaires sont implémentées : les réseaux de neurones fournissent un cadre basé sur le modèle connexionniste du cerveau (apprentissage approfondi, *deep learning*) et en reproduisent certaines propriétés, qui leur permettent par exemple de mimer avec succès la reconnaissance d'objets, l'intelligence comporte aussi un processeur symbolique et logique capable de travailler à un niveau plus haut et plus abstrait de description. Ces deux technologies sont utilisées séparément ou ensemble pour reproduire les mécanismes de l'intelligence. L'induction est le mécanisme qui, à partir des données et des observations, permet d'inférer (la plausibilité d') une loi. En logique, l'induction n'apporte pas de certitude. Le raisonnement déductif au contraire, déduit des faits nouveaux à partir de faits élémentaires et vérifie des propriétés par inférence logique. La conclusion est aussi valide que les prémisses. Les logiques modales ou la gestion des probabilités ajoutée aux logiques de descriptions permet de raisonner aussi en incertitude. Le paradigme logique sous-jacent revêt un très grand intérêt pour la vérification de la cohérence d'un système qui prend en compte de plusieurs tâches, pour

aboutir par exemple à une décision clinique. L'abduction est une autre forme d'inférence qui, partant d'une observation ou d'un ensemble d'observations, recherche l'explication la plus simple ou la plus probable aux phénomènes observés. Elle sert par exemple à générer des hypothèses diagnostiques. C'est aussi le mécanisme des systèmes experts en chaînage arrière (orientés par un but) qui recherchent les règles d'inférence qui ont une partie droite qui correspond au but et déclenchent dans ce cas la recherche de preuve des parties gauches de la règle. De nombreux scores sont associés à l'induction, valeurs prédictives, scores de vraisemblance, et l'inférence abductive est étroitement liée à la prédiction. Ces différents mécanismes s'articulent dans l'intelligence artificielle.

2. Approches inductives et connexionnistes en radio-oncologie

Les modèles récents d'apprentissage sont basés sur des approches dites connexionnistes à base de réseaux de neurones. En médecine l'apprentissage approfondi requiert de gros volumes de données pour atteindre de bonnes performances. La collecte massive de données en vie réelle, par les systèmes d'information utilisés pour le soin, et leur intégration dans un entrepôt de données est indispensable à la construction de ces algorithmes. De nombreux hôpitaux universitaires ont développé dans le courant des années 2000 des entrepôts de données cliniques centralisant les données des patients, la plupart selon le standard i2b2, développé à Harvard. Cependant le fait que les logiciels spécifiques à la radiothérapie aient souvent fonctionné « en *stand-alone* » a différé l'intégration de ces données dans les entrepôts de données hospitaliers. Dans l'objectif de disposer d'une intelligence artificielle qui ne fonctionne pas en silo mais qui prenne en compte tous les paramètres entrant dans la médecine de précision, les entrepôts doivent contenir toutes les informations relatives à la tumeur et au patient. L'hôpital européen Georges-Pompidou, établissement de l'Assistance publique-hôpitaux de Paris (AP-HP), a mis en place en 2008, un entrepôt sur le modèle i2b2 pour servir les besoins en termes de recherche, et a été le premier à y intégrer les données de radiothérapie (1,2).

Ces modèles d'intelligence artificielle ont obtenu leurs meilleures performances sur des données dites « non structurées » à savoir images et textes. En ce qui concerne les images, la communauté de recherche de *computer vision* est extrêmement active sur la classification des photos d'individus sur les réseaux sociaux, l'identification de sujets dans les images dynamiques, et la reconnaissance d'objets dans des images complexes. Ces mécanismes sont très proches de la classification des images médicales, comme cela a été montré par la comparaison des tâches de classification des lésions radiologiques et des dessins d'animaux qui activent les mêmes zones cérébrales et sont réalisées par des personnes entraînées en environ 1,3 s toutes les deux (3). La classification automatique de lésions cutanées par des réseaux de neurones offre des performances similaires à la classification de ces lésions par des dermatologues (4). De même, l'analyse des lames en anatomie pathologique bénéficie depuis plusieurs années d'approches d'apprentissage approfondi pour la détection des mitoses, la segmentation et la classification en cancer ou en lésion bénigne (5). Les réseaux de neurones sont

aussi utilisés pour la classification des images radiologiques, par exemple des nodules pulmonaires, avec des résultats encourageants (6,7).

Le point de vigilance reste la question de la sélection des caractéristiques (*features*) de l'image, traditionnellement inspirée par la littérature et des avis d'experts. En 2012, Lambin et al. ont introduit le terme de radiomique pour décrire l'extraction à haut débit de caractéristiques quantitatives des images en vue de les utiliser pour l'aide au diagnostic, au pronostic, et au suivi de la réponse au traitement (8). La radiomique peut être divisée en plusieurs étapes, chacune avec ses propres défis : (i) l'acquisition et la reconstruction d'image, (ii) la segmentation, (iii) l'extraction des paramètres et la quantification (iv) les bases de données intégratives et le partage, (v) le processus d'analyse de données multidimensionnelles (9). Suite à cette publication, la plupart des algorithmes ont comporté une phase séparée d'extraction des paramètres (*features*) puis la classification des images proprement dite (10), alors que, grâce aux travaux fondamentaux de l'intelligence artificielle, on s'oriente maintenant vers l'utilisation des réseaux de neurones directement sur des images brutes sans étape préalable de radiomique pour extraire les paramètres.

En plus des images, le phénotypage haut débit doit exploiter les données textuelles. Le phénotypage vise non pas seulement à extraire pour une cohorte d'individus uniquement un nombre limité de caractéristiques identifiées dans la littérature ou par les experts comme importantes mais à extraire toutes les caractéristiques phénotypiques pour tous les patients à partir de toutes les données du dossier patient. On parle aussi de *deep phenotyping*. Deux systèmes illustrent bien cette approche, Dr Warehouse® développé initialement à l'institut Imagine et à l'hôpital Necker pour phénotyper les patients atteints de maladies rares, et DeepPhe développé à Harvard pour phénotyper les patients atteints de cancer (11,12). Dr Warehouse® dispose d'un module qui extrait tous les faits rapportés dans le dossier médical du patient, en gérant les négations, les histoires familiales et les ambiguïtés dans les textes et les représente dans un format unifié basé sur le Unified Medical Language System (UMLS) (13). De manière similaire pour l'anglais, DeepPhe comporte un flux de traitement des données textuelles qui aboutit à un phénotypage *all for all*.

3. Approches déductives et logiques en radio-oncologie

En complément des approches connexionnistes, l'intelligence artificielle est aussi fondée sur la modélisation *top-down* du raisonnement logique, la représentation et la manipulation de connaissances par des symboles et des logiques formelles. En oncologie, les modèles dits mécanistiques du cancer explicitent les connaissances fondamentales sur les processus et l'histoire de la maladie indispensables à l'« explicabilité » de l'intelligence artificielle, à la garantie logique de l'intelligence artificielle dans la résolution de problèmes, et à l'exécution de tâches complexes de diagnostic et de traitement.

Les ontologies fournissent un ensemble standardisé de concepts et de relations entre concepts qui représentent un domaine et garantissent l'interopérabilité des bases de données. De nombreuses ontologies existent dans le domaine biomédical, implémentées en logiques de description (par exemple, OWL), qui apportent des capacités de raisonnement sur les classes d'objets et les instances qui représentent les individus. L'anatomie par exemple bénéficie d'une ontologie formelle de grande envergure, le Foundational Model of Anatomy (FMA), utilisé par de nombreuses applications médicales (14). Cependant en radiothérapie, des adaptations spécifiques sont nécessaires. La Radiation Oncology Structures a été développée dans ce but et va être utilisée pour intégrer les données dosimétriques au niveau de l'AP-HP, soit plus de 7 millions de patients (en 2019) (15). Cette ontologie complète d'autres initiatives en radiothérapie comme la Radiation Oncology Ontology (ROO) (16).

Les approches logiques permettent aussi d'implémenter des règles de décision établies (recommandations internationales), pour guider la décision médicale au cours des réunions de concertation pluridisciplinaires et de tester la cohérence des règles de manière automatique, comme cela a été fait pour le cancer du sein (17,18).

4. Approches abductives et prédiction en radio-oncologie

Prédiction et pronostic sont des concepts liés, (*cancer prognostication*) et sont souvent ramenés à des tâches de stratification et de classification. Les réseaux de neurones peuvent être utilisés et ont obtenu des résultats intéressants sur plusieurs cohortes rétrospectives ; par exemple pour le cancer du poumon (19) ; un résultat intéressant est la présence de signes phénotypiques corrélés au cycle cellulaire et à la transcription.

Un enjeu majeur est la prédiction de la réponse à un type de chimiothérapie, de manière à choisir le médicament le plus adapté ; à ce jour la plupart des modèles utilisés dans le cadre de la recherche translationnelle et du soin sont mécanistiques et implémentés sur des plateformes comme celle d'OncoTrack qui intègrent des données de cohortes, des modèles animaux et des modèles cellulaires (20,21). Des réseaux de neurones ont aussi obtenu des résultats encourageants pour prédire la présence du biomarqueur MSI à partir des lames d'anatomie pathologique de cancers digestifs pour décider d'une immunothérapie et prédire avec la radiomique la réponse à la radiothérapie des cancers du rectum (22-24).

5. Discussion

Le paysage informatique de l'intelligence artificielle est centré sur les approches neuronales et logiques. Ces technologies sont déjà largement utilisées en radiothérapie, en particulier les réseaux de neurones (25). D'autres auteurs utilisent le terme d'intelligence artificielle pour des méthodes statistiques, que nous n'abordons pas mais qui peuvent être suffisantes pour certaines tâches. Ici, nous nous sommes concentrés sur les formes de l'intelligence artificielle, à savoir raisonnement logique,

reconnaissance des formes et perception, traitement du langage naturel et planning/navigation. Cela appelle plusieurs réflexions.

Lorsque l'intelligence artificielle cherche à résoudre des problèmes médicaux complexes, elle devient nécessairement hybride et combine l'apprentissage approfondi et le raisonnement. Aujourd'hui les applications sont de type « étroit » (*artificial narrow intelligence*), ne s'intéressant qu'à un problème bien défini, sans véritable capacité de s'adapter aux situations imprévues. Des applications plus générales sont en préparation, comme l'aide à la décision médicale dans le cadre la médecine de précision (26) qui requiert modélisation biologique et apprentissage continu, le modèle à un temps t étant challengé par les données qui s'accumulent. En outre, l'oncologie est un des domaines des projets de jumeaux numériques (*digital twins*). Ceux-ci correspondent à la modélisation d'un objet physique ou d'un individu dans le monde virtuel. Les avantages sont de reproduire l'évolution physique sur l'objet virtuel pour appréhender les phénomènes, et de pouvoir faire des tests sur le jumeau virtuel pour prédire et analyser la réponse au traitement. Le jumeau numérique permet également un partage d'informations entre acteurs de la prise en charge.

En plus de se rapprocher de l'intelligence humaine, l'intelligence artificielle permet de résoudre des problèmes tels que l'adaptation du traitement en radiothérapie pour un même patient (en raison de la réduction de volume de la tumeur, de la perte de poids, etc.) avec des résultats prometteurs pour les réseaux de neurones. Ces tâches bénéficient des avancées en termes d'apprentissage approfondi, comme les réseaux antagonistes génératifs (*generative adversarial networks*, GAN), considérés par Yann Le Cun comme 'l'idée la plus intéressante depuis 10 ans ». L'idée repose sur la combinaison de générateurs et de discriminateurs qui en quelque sorte « dialoguent » pour réussir à générer des données synthétiques fiables (27).

Les algorithmes d'intelligence artificielle deviennent des dispositifs médicaux et seront évalués comme tels. Là où les algorithmes déterministes peuvent être facilement validés, et leur cohérence contrôlée, les autres dispositifs doivent fournir des preuves de leur qualité et faire l'objet de validation externe. Pour cette validation, les initiatives telles que le Health Data Hub permettront de les tester sur des jeux de données extérieurs indépendants. Ces contraintes sont absolument nécessaires pour en garantir la sécurité.

Pas de conflit d'intérêts

Références

1. Jannot A-S, Zapletal E, Avillach P, Mamzer M-F, Burgun A, Degoulet P. The Georges-Pompidou University Hospital Clinical Data Warehouse: A 8-years follow-up experience. *Int J Med Inf* 2017;102:21-8.

2. Zapletal E, Bibault JE, Giraud P, Burgun A. Integrating multimodal radiation therapy data into i2b2. *Appl Clin Inform* 2018 Apr;9(2):377-390. doi:10.1055/s-0038-1651497.
3. Melo M, Scarpin DJ, Amaro E Jr, Passos RB, Sato JR, Friston KJ et al.. How doctors generate diagnostic hypotheses: a study of radiological diagnosis with functional magnetic resonance imaging. *PLoS One* 2011;6(12):e28752. doi: 10.1371/journal.pone.0028752.
4. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115-118. doi: 10.1038/nature21056.
5. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J Pathol Inform* 2016;7:29. doi: 10.4103/2153-3539.186902.
6. Ather S, Kadir T, Gleeson F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. *Clin Radiol.* 2019; doi: 10.1016/j.crad.2019.04.017.
7. Jung H, Kim B, Lee I, Lee J, Kang J. Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method. *BMC Med Imaging* 2018;18(1):48. doi:10.1186/s12880-018-0286-0.
8. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48(4):441-6. doi: 10.1016/j.ejca.2011.11.036.
9. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30(9):1234-48. doi: 10.1016/j.mri.2012.06.010.
10. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015;5:13087. doi: 10.1038/srep13087.
11. Garcelon N, Neuraz A, Salomon R, Faour H, Benoît V, Delapalme A et al. A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse. *J Biomed Inform.* 2018;80:52-63. doi: 10.1016/j.jbi.2018.02.019.
12. Savova GK, Tseytlin E, Finan S, Castine M, Miller T, Medvedeva O, et al. DeepPhe: A natural language processing system for extracting cancer phenotypes from clinical records. *Cancer Res* 2017;77(21):e115-e118. doi: 10.1158/0008-5472.CAN-17-0615.

13. Garcelon N, Neuraz A, Salomon R, Bahi-Buisson N, Amiel J, Picard C, et al. Next generation phenotyping using narrative reports in a rare disease clinical data warehouse. *Orphanet J Rare Dis* 2018;13(1):85. doi: 10.1186/s13023-018-0830-6.
14. Detwiler LT, Mejino JLV, Brinkley JF. From frames to OWL2: Converting the foundational model of anatomy. *Artif Intell Med* 2016;69:12-21.
15. Bibault JE, Zapletal E, Rance B, Giraud P, Burgun A. Labeling for Big Data in radiation oncology: The Radiation Oncology Structures ontology. *PLoS One* 2018;13(1):e0191263. doi: 10.1371/journal.pone.0191263.
16. Traverso A, van Soest J, Wee L, Dekker A. The radiation oncology ontology (ROO): Publishing linked data in radiation oncology using semantic web and ontology techniques. *Med Phys* 2018;45(10):e854-e862. doi: 10.1002/mp.12879.
17. Séroussi B, Prebet C, Ngo C, Teixeira L, Guézennec G, Bouaud J. Using the recommendations inferred by a decision support system to compare breast cancer clinical practice guidelines. *Stud Health Technol Inform* 2019;258:60-64.
18. Séroussi B, Guézennec G, Lamy JB, Muro N, Larburu N, Sekar B et al. Reconciliation of multiple guidelines for decision support: a case study on the multidisciplinary management of breast cancer within the DESIREE project. *AMIA Annu Symp Proc* 2018;2017:1527-1536.
19. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLoS Med* 2018;15(11):e1002711. doi: 10.1371/journal.pmed.1002711.
20. Fröhlich F, Kessler T, Weindl D, Shadrin A, Schmiester L, Hache H et al. Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. *Cell Syst* 2018;7(6):567-579.e6. doi: 10.1016/j.cels.2018.10.013.
21. Gu W, Yildirimman R, Van der Stuyft E, Verbeeck D, Herzinger S, Satagopam V, et al. Data and knowledge management in translational research: implementation of the eTRIKS platform for the IMI OncoTrack consortium. *BMC Bioinformatics* 2019;20(1):164. doi: 10.1186/s12859-019-2748-y.
22. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH et al.. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019;25(7):1054-1056. doi: 10.1038/s41591-019-0462-y.
23. Peecken JC, Nüsslin F, Combs SE. "Radio-oncomics" : The potential of radiomics in radiation oncology. *Strahlenther Onkol* 2017;193(10):767-779. doi:10.1007/s00066-017-1175-0

24. Bibault JE, Giraud P, Housset M, Durdux C, Taieb J, Berger A et al. Deep learning and radiomics predict complete response after neoadjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep* 2018;8(1):12611. doi: 10.1038/s41598-018-30657-6. Erratum in: *Sci Rep* 2018;8(1):16914.
25. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med* 2018;98:126-146. doi: 10.1016/j.combiomed.2018.05.018.
26. Parekh VS, Jacobs MA. Deep learning and radiomics in precision medicine. *Expert Rev Precis Med Drug Dev* 2019;4(2):59-72. doi: 10.1080/23808993.2019.1585805.
27. Liang X, Chen L, Nguyen D, Zhou Z, Gu X, Yang M, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys Med Biol* 2019;64(12):125002. doi: 10.1088/1361-6560/ab22f9.