

L'imagerie médicale à l'heure de l'intelligence artificielle

Nicholas Ayache

Membre de l'Académie des sciences
Inria, Directeur de recherche,
2004 Route des Lucioles
06902 Sophia Antipolis
France

L'intelligence artificielle bouleverse le monde de l'imagerie médicale. Les algorithmes et les modèles, déjà très présents dans la construction même des images médicales, sont aujourd'hui suffisamment puissants pour guider l'analyse des images médicales aussi bien, voire mieux, que des experts humains. Ces algorithmes permettent de construire des représentations numériques du patient qui guident le diagnostic, le pronostic et la prise en charge thérapeutique. Le 2 mai 2018 un colloque au Collège de France a rassemblé de nombreux experts sur ce sujet. Leurs interventions sont visibles sur le site du Collège (1).

Des exemples récents ont frappé les esprits (2). En dermatologie, un réseau convolutionnel profond, préalablement entraîné sur plus d'un million d'images naturelles, a été ensuite ajusté sur les images de 130 000 lésions dermatologiques, pour apprendre à distinguer automatiquement les lésions cancéreuses des lésions bénignes comme un dermatologue expert. En radiologie, l'entreprise française Therapixel a entraîné un réseau convolutionnel profond sur les 640 000 mammographies d'un challenge mondial, et a remporté la compétition en distinguant les mammographies suspectes des normales mieux que tous ses concurrents. Son logiciel surpasse aujourd'hui les performances des radiologues experts. Enfin, en ophtalmologie, d'autres réseaux convolutionnels profonds, entraînés sur plus de 130 000 images de la rétine, sont capables de détecter les rétinopathies diabétiques aussi bien qu'un ophtalmologue expert. Pour la première fois dans l'histoire, un logiciel (IDx-DR) a été approuvé par la Food and Drug Administration (FDA) pour faire ce diagnostic automatiquement sans que l'image soit vue par un ophtalmologue.

Même si ces résultats sont spectaculaires, ils restent pour l'instant confinés à des tâches précises, relativement étroites, sur lesquelles il existe d'immenses bases de données préalablement étiquetées par des experts. Or ces bases de données ne sont pas légion, leur acquisition et surtout leur étiquetage par des experts coûte très cher. Il faut s'assurer de la bonne représentativité des cas rares (qui par définition ne sont pas nombreux !), de l'absence de biais (genre, âge, ethnicité, etc.), de la présence de sujets sains en nombre suffisant (ce qui pose un problème éthique lorsque la modalité d'imagerie est irradiante par exemple). Recueillir de telles bases de données et les mettre à la disposition des algorithmes d'apprentissage pose également de sérieux problèmes de confidentialité.

Par ailleurs, l'intelligence artificielle ne se limite pas au développement de réseaux convolutionnels profonds, dont les millions de paramètres sont difficiles à interpréter, et dont l'ajustement requiert l'entraînement sur d'immenses bases de données. Il existe toute une variété d'algorithmes de classification qui peuvent être entraînés sur des bases de données de tailles très variables avec de très bons résultats (3). Il est également possible de construire des modèles numériques du patient qui exploitent toutes les connaissances que nous avons en anatomie et en physiologie pour se limiter à un nombre réduit de paramètres biophysiques plus facilement interprétables. Enfin, le modèle du patient numérique permet de synthétiser des images médicales suffisamment réalistes pour entraîner des systèmes d'apprentissage automatique, puis de transférer l'apprentissage à des images réelles.

Les méthodes modernes de sciences des données permettent aussi de prendre en compte des données de plus en plus « holistiques » sur le patient, incluant des données d'imagerie (structurelle et fonctionnelle), mais également des données biologiques (génétiques,

métabolomiques, etc.), voire comportementales et environnementales (style de vie, etc.). Ainsi la base de données UK-Biobank prévoit de mettre à disposition de la communauté scientifique l'ensemble de ces données acquises sur plus de 100 000 participants (4). En France, un important projet de collecte des données images associées à d'autres données biomédicales acquises sur un très grand nombre de patients est mis en place à l'APHP pour servir la communauté scientifique (5).

Le champ d'application de l'intelligence artificielle en imagerie médicale est amené à servir toutes les disciplines médicales. Sans être exhaustif, on peut déjà citer radiologie, dermatologie, ophtalmologie, neurologie, psychiatrie, cardiologie, sénologie, hépatologie, endoscopie, endomicroscopie, anatomopathologie, radiothérapie et chirurgie guidées par l'image (6, 7, 8, 9). De nouveaux problèmes de droit sont soulevés par l'apparition de ces nouveaux logiciels d'aide à la décision (10), et des questions sont posées sur l'avenir même de certaines spécialités médicales (11). L'ensemble de ces applications et de ces questions est également abordé dans de nombreux chapitres de cet ouvrage.

Pour conclure, je voudrais insister sur le fait que intelligence artificielle et patient numérique sont un ensemble d'outils informatiques au service d'une médecine aux « 4 p » : plus **personnalisée**, **précise**, **prédictive** et **préventive** destinée à mieux soigner le patient réel. Ces nouveaux outils ont vocation à aider le médecin, pas à le remplacer. En effet certaines des qualités du médecin comme la **compassion**, la **compréhension**, l'**esprit critique** et la **conscience professionnelle** (les « 4 c ») restent encore et pour longtemps l'apanage de l'intelligence humaine. Elles demeurent aujourd'hui irremplaçables.

(1) Gérard Berry , « Imagerie médicale et apprentissage automatique : vers une intelligence artificielle ? » : <https://www.college-de-france.fr/site/gerard-berry/symposium-2017-2018.htm>

(2) Nicholas Ayache, « Patient numérique et intelligence artificielle », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-09h05.htm>

(3) Bertrand Thirion, « Introduction aux approches *Big Data* pour l'imagerie médicale », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-09h35.htm>

(4) Tom Nichols, « Challenges and Opportunities in Population Neuroimaging », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-11h30.htm>

(5) Raphaël Beaufret, « Le projet *Entrepôt de données de santé – imagerie* à l'AP-HP », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-11h00.htm>

(6) Anne Bertrand, Stanley Durrleman et Stéphane Epelbaum, « Neuroimagerie, neurologie et modèles numériques pour la maladie d'Alzheimer », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-14h30.htm>

(7) Isabelle Thomassin-Naggara et Olivier Clatz, « Technologies innovantes au chevet du dépistage du cancer du sein : quel futur pour les femmes ? », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-15h15.htm>

(8) Laure Fournier, « Applications radiomiques en cancérologie » Collège de France, 2 mai 2018 : https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-16h30_1.htm

(9) Michael Brady, « Quelques exemples de transferts du laboratoire vers la pratique », Collège de France, 2 mai 2018 : https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-17h30_1.htm

(10) Didier Truchet et Bruno Potier de la Varde, « Les questions du juriste : quelles protections pour les patients et les données ? », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-12h00.htm>

(11) Francis Besse, « Pourquoi les radiologues espèrent et redoutent à la fois l'arrivée de l'IA », Collège de France, 2 mai 2018 : <https://www.college-de-france.fr/site/gerard-berry/symposium-2018-05-02-18h00.htm>