

Le cancer du poumon ne serait pas le cancer le plus létal au Monde si son diagnostic n'était pas aussi souvent trop tardif. En effet, le cancer du poumon est très bien curable quand celui-ci est pris en charge à un stade précoce. Cependant, celui-ci s'avère silencieux sur le plan symptomatologique jusqu'à un stade trop avancé. Son dépistage précoce s'annonce alors essentiel. Celui-ci repose sur le scanner thoracique qui nécessite une longue analyse, très dépendante des connaissances et acquis des radiologues. C'est à ce niveau qu'intervient le deep learning. Ce concept correspond à un procédé d'apprentissage automatique qui utilise des neurones avec plusieurs couches de neurones cachées. Ces réseaux de neurones sont capables, dans le cadre de l'imagerie médicale, d'identifier des anomalies que l'œil du radiologue n'aurait pas vu ou encore des anomalies invisibles pour la vision humaine. C'est dans ce contexte que nous pouvons nous demander si l'intelligence artificielle peut-elle devenir un véritable outil de diagnostic du cancer du poumon et remplacer le radiologue à l'avenir?

Nous verrons dans un premier temps que l'IA est un outil qui est devenu très performant, puis dans un second temps nous étudierons les méthodes qui ont fait que l'IA est devenue si performante dans le diagnostic par imagerie. Enfin, nous verrons que cependant l'IA présente des limites.

I/L'IA, un outil très performant

1- Les problèmes que l'on cherche à résoudre avec l'IA

Dans le cadre de la lecture d'imagerie, l'IA peut présenter des avantages et des compétences qui ne seront certainement jamais atteignables par le radiologue. En effet, le deep learning se présente comme une réponse à un besoin médical de taille qui est d'améliorer la détection précoce des cancers tout en réduisant les erreurs diagnostiques. Le premier problème majeur est celui du temps entre le début de la maladie et le diagnostic, qui est de nos jours bien trop long et qui explique le caractère aussi létal du cancer du poumon. En effet réduire le temps du diagnostic serait une avancée majeure pour cette pathologie puisque l'on sait que traiter un patient dès les premiers stades de son cancer est pour lui un facteur de bon pronostic. L'objectif du deep learning est donc ici d'aider le radiologue à trouver les lésions les plus précoces sur les imageries, qui sont aussi les plus compliquées à voir pour prendre en charge le patient au plus vite. De plus, les radiologues doivent prêter une extrême attention à chaque imagerie analysée, ce qui est pourvoyeur de fatigue visuelle, de baisse de l'attention ce qui peut être à l'origine d'erreurs d'interprétation non sans conséquence. Sur ce point l'intelligence artificielle se présente comme un bel atout pour assister le radiologue par l'analyse automatique de l'ensemble des imageries en signalant les lésions suspectes. Parmi les limites majeures du dépistage actuel du cancer du poumon, on retrouve les faux positifs et les faux négatifs. En effet, des erreurs de résultats entraînent des examens complémentaires rarement appréciés par le patient ainsi que des coûts supplémentaires pour le système de santé et souvent une anxiété majeure pour les faux positifs. Concernant les faux négatifs, cela entraîne des retards sur la prise en charge qui altère fortement le pronostic du patient. C'est dans cette optique que des modèles de deep learning ont été conçus pour réduire les erreurs d'interprétation. A cela s'ajoutent les difficultés de travail des radiologues car le nombre d'examens, du fait du développement des dépistages par scanner, augmente chaque année, ce qui entraîne un manque de personnel. L'IA peut aider aussi à ce niveau en réalisant un premier tri automatique des examens afin que les radiologues puissent davantage se concentrer sur les cas les plus complexes. Enfin,

les méthodes d'analyse des radiologues reposent souvent sur des caractères bien définies tels que la taille, la forme, la densité des lésions. Néanmoins, cela expose à de nombreuses limites car cela reposait en grande partie sur les connaissances du concepteur du système, laissant passer certaines caractéristiques plus subtiles. Le deep learning se présente encore comme avantageux sur ce point puisqu'il peut apprendre automatiquement les caractéristiques à partir des images. Ainsi, l'IA permet de résoudre de nombreux problèmes liés pour certains à la nature humaine du radiologue, le tout permettant de réduire le temps du diagnostic, les erreurs d'interprétations, diminution des charges de travail. Tous ces arguments justifient l'importance du deep learning dans le diagnostic de cancer du poumon.

2- Évolutions de l'IA

Les revues étudiées illustrent bien l'évolution des différentes méthodes de l'IA dans le diagnostic du cancer du poumon. Au début, les premiers travaux étaient basés sur le machine learning, qui nécessitait une extraction manuelle des caractéristiques sur les imageries étudiées. Aujourd'hui, le machine learning a été remplacé par le deep learning qui est capable d'un apprentissage autonome des informations pertinentes directement à partir d'un scanner thoracique. Cette transition a permis une amélioration nette des performances des systèmes d'aides au diagnostic. Les Convolutional Neural Networks ont été une véritable avancée technologique dans l'analyse des images médicales. En effet, les CNN sont capables d'identifier eux-mêmes les différents éléments comme les contours, textures, formes ou irrégularités au scanner pulmonaire. Puis leurs couches plus profondes peuvent détecter les structures telles que les nodules ou anomalies évoquant une tumeur. Le CNN a ensuite été amélioré via des modèles comme VGG, ResNet, DenseNet et MobileNet qui ont chacun permis d'apporter une amélioration spécifique. Ainsi, ResNet permet de s'entraîner des réseaux profonds sans disparition du gradient, DenseNet a permis une amélioration de la circulation de l'information entre les couches, et MobileNet permet une architecture plus légère pour les applications nécessitant une faible puissance de calcul. De plus, les architectures Faster R-CNN ont permis la détection automatique des nodules. Celle-ci permet une identification rapide des zones potentiellement cancéreuses en limitant le nombre de faux positifs. Pour ce qui est de la segmentation des nodules, c'est le modèle U-Net qui permet l'évaluation de la taille du volume ou encore de la forme des lésions. De nos jours, certaines architectures émergentes inspirées des traitements de langages automatiques, les Vision Transformers, permettent une analyse simultanée dans différentes régions du scanner, ce qui permet une compréhension plus globale et la détection, parfois, d'anomalies plus discrètes. Cependant les CNN restent les références dans le traitement des informations locales, c'est pourquoi ont été créés des modèles hybrides combinant CNN et Vision Transformers qui permettent d'obtenir des résultats plus solides en termes de précision. Ainsi, le CNN-SVM est un modèle hybride qui associe un CNN et un Support Vector Machine pour améliorer la classification. Le CNN a toujours pour rôle l'extraction automatique des caractéristiques des scanners, tandis que le SVM, en remplaçant une couche du réseau CNN, permet de rendre la frontière de décision plus critique. Cependant, les dernières recherches ne cherchent plus seulement la performance et la puissance de l'algorithme mais plutôt à rendre les modèles plus interprétables et plus généralisables comme avec les approches sur l'Explainable AI, les mécanismes d'attention, les réseaux tridimensionnels, le Transfer Learning et le Federated Learning. Ainsi, l'IA, par les différentes architectures et modèles développés, est un outil qui devient de plus en plus solide dans le

domaine du diagnostic, notamment dans le cas du cancer du poumon, malgré encore quelques failles.

3- Les performances

De nos jours, le deep learning dans la lecture, l'imagerie et ici de scanner du poumon atteint des compétences allant de égales à supérieure de celle d'un radiologue très expérimenté sur des tâches très précises, notamment grâce au développement des différentes architectures de plus en plus performantes. Pour évaluer ces modèles émergents, plusieurs indicateurs sont utilisés. Premièrement, l'accuracy mesure la proportion de diagnostics correct, la sensibilité, qui évalue la capacité à reconnaître les cancers quand il y a réellement un cancer, la spécificité, qui évalue la capacité à reconnaître un patient sain lorsque celui-ci est vraiment sain, le F1-score qui équilibre précision et sensibilité. A cela, s'ajoute l'AUC, pour Area Under the Curve qui permet la comparaison entre les différents modèles notamment face à des bases de données déséquilibrées. Ces indices ont permis de prouver l'amélioration continue des architectures. En effet, des premiers réseaux de neurones convolutifs aux dernières architectures telles que ResNet, DenseNet, EfficientNet, Vision Transformers, et les modèles hybrides, les performances de l'IA n'ont fait qu'accroître au point que certaines études estiment que des modèles hybrides peuvent atteindre un taux de précision de 99%. Là où l'IA se montre d'autant plus intéressante, c'est sur la détection des nodules précoces. En effet, celle-ci est capable de détecter des nodules de parfois quelques millimètres, très difficilement visible pour l'œil humain. Cette avancée se montre d'autant plus pertinente dans le cadre du cancer du poumon étant donné que cette pathologie est très bien curable lorsque celle-ci est prise en charge précocement. Ainsi, l'IA permet de prendre de l'avance sur la prise en charge de cette maladie qui est trop souvent trop tardive, ce qui la classe à la place du cancer le plus létal. De plus, les progrès de l'IA permettent de réduire de manière significative le nombre de faux positifs sans augmenter le nombre de faux négatifs, ce qui est très intéressant aussi bien sur l'aspect économique que l'aspect psychologique pour le patient. En effet, un résultat positif pour un patient qui ne l'a réellement pas, entraîne de nombreux examens complémentaires coûteux, et l'annonce de la suspicion de cancer à un patient à une forte atteinte émotionnelle pour ce dernier qui n'est pas à négliger. Certaines revues viennent à comparer les performances de l'IA par rapport à celle d'un radiologue expérimenté. Sur des tâches bien définies, telles que la détection des nodules pulmonaires, les meilleurs modèles se montrent autant voir supérieurs au spécialiste seul. Cette conclusion reste tout de même à prendre avec précaution étant donné que les IA sont surtout meilleures, pour le moment, sur des bases de données bien définies et préparées alors que les radiologues prennent en compte le contexte clinique du patient, pas toujours disponible pour l'algorithme. Si l'on prend pour exemple, le modèle hybride CNN-SVM, celui possède une précision comprise entre 81% et 90% et une sensibilité de 85% à 95% et un F1-score pouvant aller jusqu'à 92%, et ces performances sont en constantes progressions grâce au renouvellement continu des nouveaux modèles. Sur le point des performances, l'IA est donc d'une très grande importance dans le cadre du diagnostic du cancer du poumon. Les derniers modèles sont une assistance de taille pour les radiologues. Cependant il est important de ne pas oublier les circonstances dans lesquelles les modèles ont été testés.

II/METHODES

1-Early-Stage Lung Cancer Detection Using Deep Learning Algorithms

La thèse intitulée Early-Stage Lung Cancer Detection Using Deep Learning Algorithms illustre la progression du deep learning en nous montrant que chaque nouveau modèle aboutit à de meilleures performances diagnostiques. En effet, les objectifs de cette étude étaient d'améliorer la sensibilité du diagnostic, de réduire les faux positifs, d'optimiser les performances, tout en restant interprétable par les médecins afin que ceux-ci puissent retrouver la lésion sur le scanner. Pour procéder, ils se sont appuyés sur quatre contributions scientifiques. La première consistait en l'amélioration du réseau VGG16, en modifiant les dernières couches convolutionnelles et en utilisant un Feature Pyramid Network, afin de détecter des nodules de différentes tailles de manière simultanée. La seconde permettait d'améliorer la précision, la vitesse d'apprentissage, la stabilité du modèle ainsi que la détection des nodules de différentes tailles. Celle-ci a nécessité le développement d'un réseau Optimized Multi-Scale CNN qui a apporté des innovations telles que l'extraction multi-échelle des caractéristiques, l'optimisation automatique des paramètres via les algorithmes Harmony Search et Beetle Antennae Search et une optimisation du réseau meilleure. La troisième contribution consistait à concentrer l'attention du réseau uniquement sur les régions du scanner qui étaient réellement importantes permettant ainsi de diminuer de manière importante le nombre de faux positifs en conservant une excellente sensibilité. Enfin, la quatrième contribution avait pour objectif de fournir une très bonne précision tout en expliquant visuellement des décisions de l'intelligence artificielle. Pour cela, il a fallu combiner CNN-3D, Swin transformer, U-Net pré entraîné et Anatomical Attention Gate. Il s'agissait du modèle le plus performant de ceux présentés de cette étude. Ces quatre modèles ont prouvé une amélioration progressive des performances de l'IA sur la détection des nodules sur la scanner. Ainsi, cela prouve qu'avec les progrès de l'IA, celle-ci devient un outil très intéressant dans l'interprétation des imageries diagnostiques.

2-CNN-SVM

L'une des plus importantes limites de deep learning dans le cadre du dépistage du cancer du poumon par scanner, est le manque de données pour avoir un entraînement réellement complet et efficace. Cependant, le CNN-SVM est un modèle hybride associant un réseau de neurones convolutif et une machine à vecteurs de support permettant d'obtenir de bonnes performances malgré la taille limitée des bases de données médicales. Pour développer ce concept, les auteurs sont partis du constat qu'une trop grande partie des scanner qui servaient d'entraînement était des imageries de patient sain et que par conséquent cela augmentait le risque d'overfitting comme avec le Softmax. Pour résoudre cette problématique, les auteurs ont donc allié le CNN, qui avait pour unique rôle de l'analyse, l'extraction automatique des caractéristiques sur l'imagerie, au SVM qui lui s'occupait de la classification finale. Ceci permettait dans un premier temps de résumer les informations essentielles de l'image. Dans un second temps, le SVM construit un hyperplan de séparation qui se divise en deux catégories: les nodules malins et les nodules bénins. De plus, les auteurs avaient mis en place des techniques de pré-traitement et de data augmentation pour que le système soit d'autant plus performant. Pour procéder à cela, les scanners avaient été normalisés selon les unités de Hounsfield, les poumons ont été segmentés pour éliminer les structures anatomiques inutiles, enfin, les images ont été transformées par des rotations, retournement, modification de la luminosité et des changements d'échelle, de translations, ce qui permet d'acquérir de manière artificielle de nouveaux exemples. Le point fort du SVM par rapport au Softmax, est que celui-ci possède une capacité de généralisation plus importante, et est moins sensible au bruit, mais l'avantage majeur de cette méthode est que l'IA est moins dépendante des bases de

données relativement pauvres. Les résultats de ce modèle étaient une précision entre 80 et 90%, sensibilité entre 85% et 95% ainsi qu'un F1-score compris entre 85% et 92%, démontrant ainsi que le remplacement du Softmax par le SVM améliore largement les performances de l'IA sur le diagnostic. Cette émergence de modèle hybride montre notamment que le deep learning dans le diagnostic ne dépend plus d'une unique méthode mais plutôt d'une combinaison de modèles pour répondre aux limites que présente l'IA. Ainsi, en combinant différents modèles, l'IA se montre comme un outil clé dans le diagnostic du cancer du poumon par scanner.

3- Orientation du choix de modèle

La revue Scientific Reports compare les différents modèles de deep learning en s'intéressant à ceux qui apportent les résultats les plus intéressants. En effet, de nombreuses architectures de deep learning ont vu le jour ces dernières années comme les CNN, les 3D-CNN, les Vision Transformers ainsi que les modèles hybrides. Chacun d'entre eux possède des avantages qui leur sont propres. Cette revue cherche alors à orienter les radiologues vers les architectures les plus pertinentes pour leur utilisation médicale. Pour cela, les auteurs ont réalisé une étude de cohorte sur des scanners thoraciques de l'un des plus importants programmes de dépistage du cancer du poumon, le National Lung Screening Trial. Cette revue comparait simultanément des modèles 3D et des modèles 2D. Du côté des modèles 2D, ceux-ci présentaient une limite importante qui était que les images étaient étudiées de manière indépendante les unes des autres, au risque de perdre des informations sur la vision tridimensionnelle du poumon. Pour les images 3D, l'avantage était, par conséquent, une meilleure compréhension de la structure anatomique du poumon. De ce fait, les images 3D devaient théoriquement se rapprocher au plus du raisonnement des radiologues quand ils examinent un scanner par leurs soins. Ceci était confirmé par les résultats qui retrouvaient une accuracy de 79%, un F1-score de 0.67 et une AUROC de 0.86, les performances des modèles 3D étaient alors bien supérieures à celles des modèles 2D. De ce fait, l'utilisation des modèles 3D permet d'améliorer la qualité des prédictions et augmente la capacité de généralisation des modèles. De plus, les auteurs se sont aussi penchés sur la question du Transfer Learning. Pour ce faire, ils ont comparé des modèles pré-entraînés sur des bases de données généralistes, comme ImageNet et Kinetics, à des modèles pré-entraînés sur des bases radiologiques spécialisées, comme RadImageNet, nnU-Net ou encore 3DSeg-8. Cette comparaison aboutit à la conclusion que les modèles pré-entraînés sur des bases de données généralistes obtiennent de meilleurs résultats. Ainsi, cette comparaison des différents modèles permet d'une part d'orienter les recherches vers un modèle qui semble plus prometteur, le modèle 3D, et permet donc d'orienter vers une encore plus importante marge de progression le deep learning dans le cadre du diagnostic du cancer du poumon.

III/LES LIMITES DE L'IA

1- Manque de base de données

L'une des limites les plus importantes du deep learning dans le diagnostic du cancer du poumon est son manque de bases de données. En effet, les réseaux neuronaux du deep learning nécessitent au minimum des dizaines de milliers pour distinguer les différentes situations possibles. Si ce manque de données est difficilement résolvable pour le moment, c'est parce qu'en médecine, le scanner est associé à une biopsie ainsi qu'un suivi clinique afin d'en arriver au diagnostic. Les bases de données référentes telles que LIDC-IDRI,

LUNA16, NLST et LNDb sont de taille bien trop modeste par rapport aux bases de données d'autres domaines. La revue systématique englobant les travaux allant de 2015 à 2024 montre que LILDL-IDRI est une base de données de référence dans le cadre du diagnostic du cancer du poumon, pourtant, celle-ci ne contient que 1000 scanners annotés, un nombre bien trop faible pour un modèle de deep learning moderne. De plus, les bases de données disponibles sont souvent déséquilibrées. En effet, celle-ci présente comme défaut d'être trop entraînées sur des imageries de poumons sains ou de cas atypiques de cancer pulmonaire plutôt que sur des cas typiques. Ce manque de diversité dans les données est appelé le *domain shift*. Pour essayer de pallier ça, plusieurs études proposent des techniques de Data Augmentation qui permettent d'augmenter la taille des bases de données en créant des images de plusieurs situations possibles. Le Transfert Learning est, dans ce contexte, très utilisé par les chercheurs. En effet, celui-ci consiste à utiliser des réseaux neuronaux entraînés sur de très grandes bases de données telles que ImageNet ou RadImageNet, avant de les spécialiser dans les cancers pulmonaires. Cette méthode permet une amélioration significative des performances selon l'étude [Scientific Reports](#). L'IA possède donc une marge de progression majeure liée à son manque de données, qui nécessite une réelle amélioration si celle-ci doit devenir un outil diagnostique nécessaire dans le cadre du cancer du poumon.

2-Des modèles peu généralisables

Actuellement, un des principaux objectifs du deep learning dans le cadre du diagnostic du cancer du poumon est de rendre les bases de données généralisables. En effet, les performances du deep learning actuel sont très dépendantes des bases de données sur lesquelles il a été entraîné au moment de son développement. Cependant, cela présente l'inconvénient que ces modèles ne seront pas aussi efficaces entre les différents hôpitaux. Effectivement, certaines caractéristiques sur les imageries entre les hôpitaux varient, comme les résolutions, les protocoles d'acquisition, les doses de rayonnement et les méthodes de reconstruction d'image. Ainsi, les reviews dénoncent que la majorité des études sont basées sur les bases de données publiques comme LIDC-IDRI, LUNA16 et NLST. Pour ces bases de données, les résultats du deep learning sont remarquables avec une sensibilité dépassant les 95%, mais ceci est expliqué par le fait que les données issues de ces bases sont très contrôlées et sont souvent ressemblantes entre elles, avec les mêmes caractéristiques. L'étude comparative [Scientific Reports](#) démontre cette limite du deep learning. Les auteurs ont évalué différents modèles d'architectures modernes dont des modèles 3D, 2D, et des Vision Transformers ainsi que différentes variantes de ResNet et MobileNet. À l'issue de cette étude, les auteurs confirment que les résultats dépendent énormément des données utilisées pour l'entraînement et du pré-entraînement. Cette étude comparative prouve que la comparaison des architectures dépend non seulement des performances de l'algorithme mais aussi de la capacité de généralisation, la généralisation étant la capacité d'un modèle à conserver ces performances lorsqu'il est face à de nouvelles données. De plus, le surapprentissage, représente aussi un obstacle, en effet, lorsqu'un algorithme est trop entraîné sur une image, celui-ci peut avoir tendance à se focaliser sur des détails trop précis, ce qui aboutit à d'excellents résultats sur les données qu'il connaît mais une moins bonne performance sur de nouvelles données. Pour résoudre ce problème, les chercheurs essaient de créer des bases de données multicentriques composées de données issues de plusieurs hôpitaux différents et de populations plus variées. Ainsi, la généralisation se montre encore comme une limite de l'intelligence artificielle dans le diagnostic du cancer du poumon.

3-La boîte noire

Une des limites majeures de l'IA est celle de la boîte noire. En effet, quand l'IA donne un résultat, le radiologue n'a pas accès à tout le raisonnement qui a eu lieu derrière pour affirmer ce diagnostic contrairement au radiologue qui est capable de justifier et donner des arguments qui l'ont amené à sa conclusion. Ce phénomène de boîte noire est expliqué en grande partie par la complexité des réseaux neuronaux composant le deep-learning. En effet, en prenant pour exemples les CNN telles que les Vision Transformer, celles-ci possèdent parfois des dizaines de millions de paramètres, ajustés de manière automatique lors des entraînements, ce qui rend leur cheminement totalement opaque pour l'humain. Par conséquent, cette obscurité de raisonnement rend la prise de décision compliquée par les médecins qui peuvent se montrer hésitants face à un résultat rendu par l'IA malgré le fait que certaines possèdent des sensibilité supérieures à 95%. Ainsi, cela pose un problème de confiance des radiologues envers l'IA car ceux-ci restent les responsables juridiques du diagnostic, et préfèrent, par conséquent, savoir pourquoi celui-ci est positif ou négatif. Face à cette limite, l'Explainable Artificial Intelligence ou XAI a été développée dans certaines études. Celle-ci permettra de mieux comprendre les résultats obtenus par l'IA en indiquant les éléments sur l'imagerie qui ont permis à l'IA d'en arriver à sa conclusion. Les Class Activation Maps sont les techniques les plus utilisées dans ce contexte, celles-ci permettent donc d'indiquer les zones du scanner qui ont poussé au résultat. La SHapley Additive exPlanations, quant à lui, a pour objectif de quantifier le poids de chaque caractéristique utilisé par le modèle d'optique de rendre le raisonnement toujours plus transparent. De plus, la thèse Early-Stage Lung Cancer Detection Using Deep Learning Algorithms, montre que les derniers modèles d'IA développés s'attardent à la fois sur les performances de celle-ci mais surtout sur la possibilité de les interpréter. En effet, il semble dorénavant essentiel de rendre le raisonnement acheminant au résultat accessible au radiologue pour que celui-ci puisse le comprendre et finalement donner le diagnostic au patient. Ainsi, malgré les excellentes performances de l'IA dans la détection de nodules, celle-ci se doit d'être moins opaque pour être pleinement intégrée dans la pratique médicale.

Nous pouvons en conclure que l'IA est un outil majeur dans le diagnostic scanner du cancer du poumon, et se présente aujourd'hui comme une clé pour réduire la mortalité du cancer du poumon qui reste le cancer le plus létal. En effet, avec les développements des différents modèles et architectures le deep learning s'est considérablement amélioré, permettant de dépasser les capacités et les compétences du radiologue. Cependant l'intelligence artificielle présente encore une marge de progression importante sur plusieurs points dont le manque de base de données, le manque de transparence du raisonnement et la difficulté à généraliser les modèles.

