



الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université Frères Mentouri Constantine
Faculté des Sciences de la Nature et de la Vie

جامعة قسنطينة 1 الإخوة منتوري
كلية علوم الطبيعة والحياة

Département : Biologie Appliquée

قسم البيولوجيا التطبيقية

Mémoire présenté en vue de l'obtention du diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Biotechnologie

Spécialité : Bio-informatique

Analyse des données cliniques et histopathologiques par des approches basées sur l'intelligence artificielle pour le diagnostic du cancer de rein

Présenté par :

Le : 25/06/2025

- Sahraoui Tahani
- Abbas Amina

Encadrant :

Jury d'évaluation:

- Dr. Bensaada Mostefa
(PROF- UFM Constantine 1)

- Presidente: Dr. Djezzar Nedjma
(PROF-UFM Constantine 1)

Co-Encadrant :

- Examinateur : Dr. Krid Adel
(PROF-UFM Constantine)

Dr. Chaita Kamel (Médecin spécialisé en
anatomopathologie)

Année universitaire
2023 – 2024

Remercîments

Nous tenons tout d'abord à adresser nos plus sincères remerciements à **Madame Djezzar Nedjma la Présidente du jury** ainsi qu'à **Monsieur Krid Adel l'Examinateur**, pour l'honneur qu'ils nous font en acceptant d'évaluer ce travail. Leur regard attentif, leur bienveillance et leurs remarques constructives sont pour nous une véritable marque de considération, et nous leur en sommes profondément reconnaissants.

J'exprime notre gratitude la plus sincère à **Dr. Bensaada**, notre encadreur, pour son expertise, la clarté de ses orientations, et la qualité de son accompagnement. Son exigence méthodologique, sa disponibilité et ses conseils avisés ont largement contribué à la concrétisation de ce mémoire.

Nos remerciements vont également à **Dr. Chaita**, notre co-encadreur, pour son implication, ses suggestions pertinentes et son soutien constant tout au long de cette étude.

Nous remercions chaleureusement **Monsieur le Chef du Département de Biologie Appliquée**, pour son appui institutionnel et son engagement à promouvoir un environnement de travail stimulant et rigoureux.

Nos remerciements s'adressent aussi à **Monsieur le Doyen**, pour son rôle essentiel dans le développement d'un cadre académique favorable à l'apprentissage, à la recherche et à l'épanouissement intellectuel.

Nous adressons également une pensée reconnaissante à l'ensemble du **personnel administratif**, pour son professionnalisme, sa réactivité et l'efficacité de son accompagnement dans toutes les démarches liées à ce travail.

Nous tenons à exprimer toute notre reconnaissance à **l'ensemble des enseignants qui ont marqué notre parcours éducatif**, depuis les premières années de ma scolarité jusqu'à l'achèvement de ce cycle universitaire.

Une reconnaissance toute particulière revient à **nos professeurs de la spécialité Bioinformatique**, qui, au cours des trois dernières années, ont su transmettre leur savoir avec passion, rigueur et générosité. Leur engagement pédagogique, leur disponibilité et la richesse de leurs enseignements ont profondément nourri notre intérêt pour cette discipline et nous' ont accompagné vers une meilleure compréhension des enjeux scientifiques contemporains.

Enfin, nous remercions toutes les personnes qui, de près ou de loin, ont contribué à la réalisation de ce mémoire. À **nos famille, à nos amis, à nos camarades** de promotion et à toutes celles et ceux dont le soutien, les encouragements ou la simple présence ont été précieux : recevez ici l'expression de notre reconnaissance la plus sincère.

Dédicace

Après la grâce et la facilitation de Dieu, qui a guidé mes pas et nous a soutenues à chaque étape de ce parcours,

Je dédie humblement ce travail à **nos honorables professeurs**, véritables flambeaux de lumière sur notre chemin. Ils ont éclairé notre route vers le savoir, éveillé en nous l'amour de la science et la passion de la recherche, et nous ont appris la rigueur et l'excellence. Leur accompagnement constant et leurs conseils précieux ont grandement contribué à l'aboutissement de ce mémoire. À eux, j'adresse toute ma reconnaissance et mon profond respect.

À **ma chère mère**, source inépuisable d'amour, de tendresse et de sacrifices. Ton soutien inconditionnel, tes prières et ta présence bienveillante ont été la force silencieuse derrière chacune de mes réussites. Que Dieu te protège et te récompense pour tout ce que tu as donné.

À la mémoire de **mon père bien-aimé**, qui a quitté ce monde mais dont les valeurs, les conseils et l'amour demeurent à jamais gravés dans mon cœur.

Je prie le Tout-Puissant de lui accorder Sa miséricorde et Son pardon, et que chaque personne qui lira ce travail élève une prière pour lui : « *Qu'Allah lui fasse miséricorde et lui accorde le pardon.* »

À **mes amis et collègues**, compagnons de route dans l'effort et la persévérance. Avec vous, les défis sont devenus des opportunités de croissance, portés par l'esprit d'entraide et la force de l'amitié sincère.

À vous tous, j'offre ce modeste travail, en témoignage d'un profond attachement, d'un amour sincère et d'une gratitude que les mots ne sauraient pleinement exprimer.

Résumé

Ce travail de recherche vise à développer un modèle d'intelligence artificielle contribuant à améliorer la rapidité du diagnostic histologique du cancer du rein, un type de cancer silencieux dont la détection précoce reste difficile. Le travail s'articule autour de l'analyse de données provenant de sources multiples, incluant des images médicales (radiologiques et histopathologiques) ainsi que des données cliniques réelles, en utilisant des techniques avancées d'apprentissage profond (Deep Learning) via le modèle ResNet50.

Les données ont été collectées à partir de la plateforme Kaggle et le service d'anatomopathologie de la clinique rénale de Daksi Constantine. Elles ont été soigneusement classées et prétraitées avant l'entraînement du modèle.

Une interface web interactive a également été développée pour permettre à l'utilisateur d'obtenir automatiquement les résultats d'analyses histologiques.

Les résultats que nous avons obtenus ont montré une grande efficacité du modèle dans la distinction entre les tissus sains et les tissus atteints, ce qui souligne la capacité de l'intelligence artificielle à soutenir la décision médicale et à réduire les délais diagnostic.

Ce projet constitue une avancée significative vers l'intégration des technologies d'intelligence artificielle dans le domaine de la santé en Algérie, en renforçant les chances de dépistage précoce et en améliorant la qualité des soins apportés aux patients.

Mots-clefs : l'intelligence artificielle, cancer de rein, Analyse des données cliniques et histopathologiqu

Abstract

This research work aims to develop an artificial intelligence model contributing to improve the speed of histological diagnosis of kidney cancer, a type of silent cancer whose early detection remains difficult. The work revolves around the analysis of data from multiple sources, including medical images (radiological and histopathological) as well as real clinical data, using advanced deep learning techniques via the ResNet50 model.

The data was collected from the Kaggle platform and the anatomical pathology department of the renal clinic of Daksi Constantine. They were carefully classified and pre-treated before the model training.

An interactive web interface has also been developed to allow the user to automatically obtain histological analysis results.

The results we obtained showed a high efficiency of the model in distinguishing between healthy tissues and affected tissues, which highlights the ability of artificial intelligence to support medical decision-making and reduce diagnostic delays.

This project is a significant step towards the integration of artificial intelligence technologies in the field of health in Algeria, by strengthening the chances of early detection and improving the quality of patient care.

Key words :artificial intelligence ,kidney cancer, Analysis of clinical and histopathological data

الملخص

يهدف هذا العمل البحثي إلى تطوير نموذج ذكاء اصطناعي يساهم في تحسين سرعة التخليص النسيجي لسرطان الكل، وهو نوع من السرطان الصامت الذي يظل اكتشافه المبكر صعبا. يدور العمل حول تحليل البيانات من مصادر متعددة، بما في ذلك الصور الطبية (الإشعاعية والنسيجية المرضية) بالإضافة إلى البيانات السريرية الحقيقية، باستخدام تقنيات التعلم العميق المتقدمة عبر نموذج ResNet50.

تم جمع البيانات من منصة Kaggle وقسم علم الأمراض التشريحي في عيادة الكل في داكسبي قسطنطين. وقد تم تصنيفها بعناية ومعالجتها مسبقاً قبل التدريب النموذجي.

كما تم تطوير واجهة ويب تفاعلية للسماح للمستخدم بالحصول على نتائج التحليل النسيجي تلقائياً. وأظهرت النتائج التي حصلنا عليها كفاءة عالية للنموذج في التمييز بين الأنسجة السليمة والأنسجة المصابة، مما يسلط الضوء على قدرة الذكاء الاصطناعي على دعم اتخاذ القرارات الطبية وتقليل التأخير في التخليص.

وبعد هذا المشروع خطوة هامة نحو دمج تقنيات الذكاء الاصطناعي في مجال الصحة في الجزائر، من خلال تعزيز فرص الكشف المبكر وتحسين جودة رعاية المرضى.

الكلمات المفتاحية: البيانات السريرية والإسثوباثولوجية، لذكاء الصناعي، سرطان الكل

Table des matières

Introduction :.....	14
1 Introduction générale :.....	14
1.1 Les stades d'evolution de la tumeur :.....	Erreur ! Signet non défini.
1.2 Les différents sous-types histologiques :	Erreur ! Signet non défini.
1.3 L'intelligence artificielle (IA) :.....	20
1.3.1 Les avantages de l'IA dans le diagnostic des maladies :.....	20
1.3.2 Les défis et limitations de l'IA :	21
2 Problématique générale :	22
3 Les Hypothèses :.....	22
Chapitre 01	2
1 Deep learning :.....	24
1.1 Méthodes et approches dans l'apprentissage profond:	25
1.2 Importance dans l'intelligence artificielle :	26
2 Machine Learning.....	26
2.1 Concepts et techniques clés.....	26
2.2 Applications du machine learning.....	27
Chapitre 02	16
1 Applications de l'IA en médecine :	29
1.1 Imagerie médicale : radiologie et histopathologie :	29
1.2 Études déjà fait sur la différenciation tumeur / non-tumeur :	30
1.3 Aide au triage et télémédecine :	31
2 Les modèles de deep learning en médecine :	31
2.1 Réseaux de neurones convolutifs (CNN) :	31
2.2 ResNet50 :	32
2.3 Autoencodeurs et GANs :	32
2.4 Transformers multimodaux :.....	32
3 Le modèle d'IA choisi pour notre étude :	33
3.1 Pourquoi :	33
3.2 Structure du ResNet50 :	35

3.2.1	Bloc résiduel :	35
3.2.2	Composition du ResNet50 :	36
	Matériels et méthode	14
1	Données utilisées :	39
2	Prétraitement des données :	40
3	Organisation et stratégie d'entraînement :	42
4	Architecture du modèle ResNet50 appliqué :	43
5	Compilation, stratégie d'optimisation et fine-tuning :	45
6	Évaluation, visualisation et test final :	45
7	Bibliothèques et environnement technique :	46
7.1	Bibliothèques utilisées :	46
7.2	Environnement d'exécution :	46
8	Interface web	47
	Résultats et discussion	14
1	Résultats radiologiques :	51
2	Résultats histopathologiques :	53
	Conclusion	14
	Bibliographie	51

Liste des figures :

Figure 1:Stade T1b. AJCC cancer staging manual. 8th edition.....	16
Figure 2:Stade T3a. AJCC cancer staging manual. 8th edition.....	17
Figure 4: Courbes ROC pour l'ensemble de test non vu en utilisant les architectures ajustées de Google Inception-V3 et ResNet50. (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018).....	34
Figure 5 : architecture du modèle ResNet50. (Mishra D. , 2020)	35
Figure 6 :Diagrammes circulaires illustrant la distribution des images radiologiques selon les sous-ensembles (Train, Validation, Test) et les classes (Normale, Tumeur).....	39
Figure 7 .Diagrammes circulaires illustrant la distribution des images histopathologies selon le grossissement (4x, 10x, 40x).	40
Figure 8 : La page d'accueil de site web	47
Figure 9 : Évolution de la précision d'entraînement et de validation au cours des époques pour le modèle ResNet50 avec fine-tuning sur un jeu de données médical binaire.	51
Figure 10: Évolution de la fonction de perte (loss) d'entraînement et de validation au cours des époques pour le modèle ResNet50 avec fine-tuning.	52
Figure 11:Courbe de validation accuracy par grossissement (résumé par moyenne des 3 folds).	53

Liste des tableaux :

Tableau 1 : Tableau récapitulatif des exemples de travaux récents en classification tumeur / non-tumeur par IA.	31
Tableau 2 : Applications principales de l'IA en médecine et outils associés. (Baker, et al., 2020) ..	31
Tableau 3 : Modèles de deep learning utilisés en médecine et leurs domaines d'application.....	33
Tableau 4: Tableau récapitulatif des exemples de travaux récents en classification tumeur / non-tumeur par IA. (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018)	33
Tableau 5 : Accuracy de validation par pli pour chaque grossissement.....	53

Introduction :

1 Introduction générale :

Le cancer du rein est une maladie maligne qui affecte les tissus à l'intérieur des reins, se caractérisant par une «masse anormale» dans les reins. Une tumeur rénale peut être bénigne, ce qui signifie qu'elle n'est pas cancéreuse et ne constitue pas une menace immédiate pour la santé. Toutefois, il est important de noter que toutes les tumeurs rénales ne sont pas cancéreuses et que tous les cancers ne partagent pas les mêmes propriétés. (Conju-Probe, 2022)

Parmi les tumeurs rénales malignes, les plus courantes sont les carcinomes, qui varient en fonction du type de cellule rénale d'où elles proviennent. Lorsqu'une tumeur maligne est suspectée, il est essentiel de distinguer son **type histologique**, c'est-à-dire de déterminer le type exact de la tumeur. Ceci est réalisé par l'examen de la tumeur, soit suite à sa résection chirurgicale, soit à l'issue d'une biopsie. (Matthew Young, carcinome à cellules rénales, 2024)

La plupart des tumeurs rénales malignes sont d'origine épithéliale ou «carcinomes» qui varient en fonction du type de cellule rénale d'où elles proviennent.

Les carcinomes rénaux non clairs représentent environ 20% de toutes les tumeurs rénales. Leur classification est continuellement mise à jour en fonction de nouvelles données cytogénétiques et moléculaires. Comme les méthodes moléculaires sont onéreuses, l'évaluation diagnostique s'appuie pour l'essentiel sur des critères morphologiques et immunohistochimiques. (Laurence Albiges 1, Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras, 2023)

La classification OMS des tumeurs rénales a nettement changé durant 50 ans. La dernière édition de l'OMS 2022 des tumeurs de l'appareil urinaire et des organes génitaux masculins prend en considération les caractères morphologiques (architecture, cytoplasme...), clinique, cytopathologique et moléculaire pour classer les tumeurs rénales.

Comme les méthodes moléculaires sont onéreuses, l'évaluation diagnostique des tumeurs rénales en pratique s'appuie pour l'essentiel sur des critères morphologiques et immunohistochimiques. (Laurence Albiges, 2023)

Selon l'examen histopathologique, le carcinome à cellules rénales claires représente près des deux tiers des tumeurs rénales.

Parmi les autres tumeurs rénales, les carcinomes rénaux d'architecture papillaire représentent le groupe le plus important et leur classification devient de plus en plus complexe. Cela inclut les carcinomes de bas grade (comme le carcinome à cellules rénales papillaire de bas grade, la tumeur papillaire à cellules claires) et les carcinomes de haut grade (comme le carcinome à cellules rénales papillaire de haut grade et les carcinomes à cellules rénales avec réarrangement TFE3 ou TFEB).

Le groupe des oncocytomes/carcinomes à cellules rénales chromophobes peut être vu comme appartenant à un même spectre lésionnel allant des néoplasies bénins (oncocytome) aux néoplasies malins (carcinome à cellules rénales chromophobes). Ces tumeurs sont parfois observées dans l'oncocytose ou le syndrome familial de Birt-Hogg-Dubé, où les cellules tumorales peuvent présenter des caractéristiques hybrides.

Les angiomyolipomes sont généralement des néoplasies bénignes mésenchymateuses, pouvant être sporadiques ou familiales (sclérose tubéreuse). Dans ce dernier cas, certains cas d'angiomyolipome épithélioïde (potentiellement malin) ont été décrits.

Les tumeurs épithéliales et stromales rénales (REST) sont un concept récent qui regroupe deux tumeurs bénignes mixtes mésenchymateuses et épithéliales : le néphrome kystique et les tumeurs mixtes épithéliales et stromales (MEST) (Mathilde Sibony 1, Carcinome rénal non à cellules claires. Mise à jour 2008 en pathologie tumorale rénale., 2008)

Mais il faut également déterminer :

- le stade d'évolution de la tumeur.
- le degré de différenciation des cellules cancéreuses (grade de Fuhrman).

1.1 Les stades d'évolution de la tumeur :

Les différents stades sont :

A. Le stade localisé : stades pT1 et pT2

La tumeur reste à l'intérieur de l'organe d'origine (le rein).

Selon la taille de la tumeur, on distingue :

- pT1a : ≤ 4 cm
- pT1b : > 4 cm jusqu'à 7 cm
- pT2a : > 7 cm jusqu'à 10 cm
- pT2b : > 10 cm

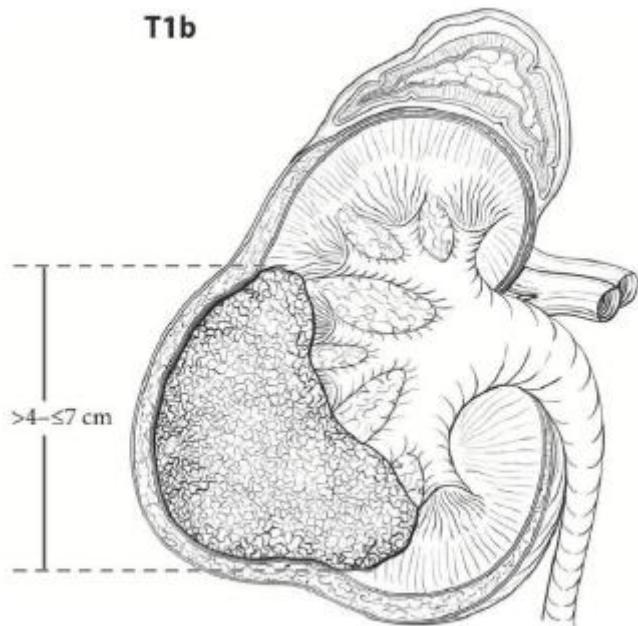


Figure 1: Stade T1b. AJCC cancer staging manual. 8th edition

B. Le stade localement avancé : Stades pT3 et pT4

La tumeur franchit l'enveloppe de cet organe (la capsule du rein) et infiltre la structure avoisinantes

- pT3a : invasion de la graisse péri-rénale, la graisse hilaire, la veine rénale ou l'une de ses branches, système pyélo-caliciel
- pT3b : extension à la veine cave sous-diaphragmatique
- pT3c : extension à la veine cave sus-diaphragmatique
- pT4 : invasion du fascia de Gerota.

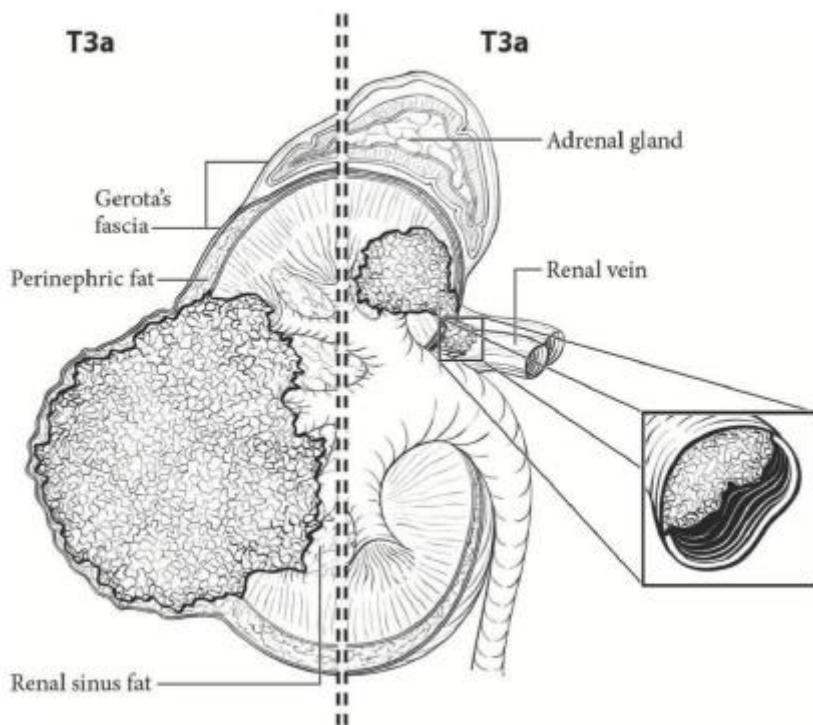


Figure 2: Stade T3a. AJCC cancer staging manual. 8th edition

C. Le stade métastatique : pN1 et pM1

La tumeur se développe à distance du rein : les ganglions lymphatiques régionaux « pN1 » ou d'autres organes « pM1 » (poumon, foie, os, cerveau, etc.).

Chaque métastase est issue de la tumeur d'origine et garde les caractéristiques de cet organe. Ainsi, des métastases au poumon sont très différentes selon qu'elles proviennent d'un cancer du rein, de la prostate ou du côlon, par exemple, justifiant des traitements potentiellement différents. (Patard, Le Cancer du rein, 2007)

La classification internationale appelée classification « pTNM » permet de caractériser les cancers selon la tumeur, les ganglions et les métastases.

La plupart des tumeurs du rein se développent dans le parenchyme rénal.

1.2 Les différents types histologiques de tumeurs rénales :

Deux groupes de tumeurs rénales se distinguent : des entités à définition morphologique (carcinome à cellules rénales claires, carcinome à cellules rénales papillaire, carcinome

à cellules rénales chromophobes, oncocytome, tumeur papillaire à cellules claires, carcinome tubulo-mucineux et à cellules fusiformes.) Et des entités à définition moléculaire (carcinome à cellules rénales avec réarrangement TFE3, avec réarrangement TFEB, ELOC muté, fumarate hydratase-déficient, succinate déshydrogénase-déficient, avec réarrangement ALK et le carcinome médullaire rénal SMARCB1-déficient)

Le carcinome à cellules rénales claires : on distingue les carcinomes à cellules rénales claires sporadiques et les carcinomes héréditaires, notamment la maladie de Von Hippel-Lindau (VHL). Dans 100 % des maladies de Von Hippel-Lindau et dans près de 70 % des carcinomes à cellules rénales claires sporadiques, il existe une mutation d'un gène appelé VHL. Ce gène joue un rôle très important dans le contrôle de l'angiogenèse. L'angiogenèse de la maladie de von Hippel-Lindau est donc devenue rapidement un objectif prioritaire dans le cancer du rein.

Le carcinome à cellules rénales papillaire : ces cancers sont le plus souvent sporadiques, mais ils peuvent être révélés sous forme héréditaire quel que soit le type. Il en existe deux sous-types selon le grade nuellaire de l'ISUP.

Le carcinome chromophobe : est une tumeur maligne, mais avec un meilleur pronostic que les deux tumeurs précédentes ; cependant, il existe des cancers chromophobes évoluant vers des métastases.

Cette classification a évolué depuis 1981. La mise à jour de 2022 les divise en six catégories selon leur morphologie ou des altérations moléculaires spécifiques. Cette classification a permis d'isoler le carcinome rénal papillaire (CRP) d'autres formes similaires caractérisées par des altérations génétiques comme les carcinomes liés aux gènes FH, TFE3 et TFEB. (Patard, Le Cancer du rein, 2007)

Désormais, il n'est plus conseillé de distinguer les CRP en types 1 et 2, mais plutôt en bas grade (ISUP 1-2) et haut grade (ISUP 3-4). Les facteurs pronostiques comme le stade pTNM, les emboles vasculaires et les composantes sarcomatoïdes restent importants, mais la nécrose n'est plus considérée comme un mauvais facteur pronostique. (Sarah, 2024)

Voici quelques études qui ont confirmé que les causes pourraient être les suivantes :

1. **Exposition aux substances chimiques nocives :** comme le trichloroéthylène, scientifiquement prouvé comme étant lié au cancer du rein après une exposition professionnelle prolongée. (Thomas Brüning 1, 2003)
2. **Facteurs génétiques :** les mutations génétiques jouent un rôle important dans certains cas de cancer du rein. (Carlo, 2022)
3. **Alimentation déséquilibrée :** la consommation d'aliments riches en graisses et de viandes transformées augmente le risque de développer différents types de tumeurs. (Yin, 2021)
4. **Le tabagisme :** constitue un facteur de risque majeur pour les cancers de l'appareil urinaire, y compris le cancer du rein. Des études ont montré que les fumeurs présentent un risque significativement accru de développer ce type de cancer, avec une relation dose-effet observée : plus la consommation de tabac est élevée, plus le risque augmente. Une méta-analyse a estimé ce risque relatif à 1,39 chez les fumeurs actuels, et jusqu'à 1,72 chez les gros fumeurs. (Xiao Qiu Liu 1, 2019)
5. **Obésité :** l'obésité est associée à une augmentation du risque de cancer du rein. (Venkatesh, 2023)
6. **Les malades du diabète :** des études épidémiologiques récentes ont mis en évidence une association significative entre le diabète de type 2 et un risque accru de développer un cancer du rein. Cette corrélation pourrait s'expliquer par l'hyperinsulinémie et l'inflammation chronique induite par le déséquilibre métabolique, favorisant ainsi la prolifération cellulaire tumorale. Une revue systématique publiée dans la revue *Clinical Diabetes* a rapporté que les patients diabétiques présentent un risque jusqu'à 40 % plus élevé de développer un cancer du rein par rapport aux non-diabétiques. Pollution environnementale : exposition prolongée à des facteurs environnementaux polluants. (Yuqing Zhang, Shufang Yang, & Zhen Wang, 2022)
7. **Facteurs immunitaires :** une faiblesse du système immunitaire peut contribuer à la formation des tumeurs.

Dans l'immense majorité des cas, il n'y a pas de signe révélateur de cancer du rein. C'est-à-dire que c'est une découverte fortuite chez un patient qui n'a pas de symptôme ; on appelle ça asymptomatique et fortuit parce qu'on le découvre à la faveur d'une échographie ou d'un scanner. Il existe un nodule sur le rein, qui est rapporté par le radiologue. La découverte de cette anomalie tissulaire, une boule de tissu au niveau du rein, doit déclencher la poursuite des explorations grâce à une consultation.

Lorsque un cancer anormal s'est propagé à d'autres organes, c'est-à-dire qu'il a donné des métastases, le traitement chirurgical peut ne plus être suffisant, et donc on a recours à des traitements généraux, qui appartiennent à deux familles : une famille qui s'appelle les anti-angiogéniques, qui sont des comprimés (qui bloquent les vaisseaux qui alimentent ces maladies).

L'immunothérapie, qui réveille votre système immunitaire, est aujourd'hui, dans le cancer du rein, ce que nous faisons : c'est que nous combinons ces deux approches ou bien nous les utilisons en séquence.

1.3 L'intelligence artificielle (IA) :

L'intelligence artificielle (IA) a marqué une avancée spectaculaire dans le domaine médical, transformant les méthodes traditionnelles en approches plus efficaces et innovantes. Cette révolution technologique permet aujourd'hui d'améliorer les processus de diagnostic, offrant des solutions personnalisées et plus précises. L'IA ne se contente plus de simplement compléter les compétences humaines, elle redéfinit fondamentalement la manière dont les diagnostics sont réalisés, rendant la médecine plus proactive et axée sur le patient.

1.3.1 Les avantages de l'IA dans le diagnostic des maladies :

1. Précision améliorée : l'IA analyse les images médicales avec une grande précision, détectant des anomalies parfois invisibles aux médecins (Chenais, 2023) (Intelligence artificielle : va-t-elle remplacer le diagnostic médical ?, 2023)
2. Diagnostic précoce : elle identifie les signes précoce de maladies graves comme le cancer, améliorant les chances de guérison. (JFR, 2024)
3. Réduction du temps de diagnostic : l'IA traite les données rapidement, accélérant le processus diagnostique. (Intelligence artificielle : va-t-elle remplacer le diagnostic médical ?, 2023)
4. Standardisation des diagnostics : elle assure une cohérence dans les résultats, réduisant les erreurs humaines. (Intelligence artificielle : va-t-elle remplacer le diagnostic médical ?, 2023)
5. Personnalisation des soins : l'IA propose des traitements adaptés en analysant les données du patient. (Imad-Addin, 2024)
6. Soutien aux professionnels de santé : elle fournit des recommandations basées sur une large base de données médicale. (Naviguer dans le paysage des soins de santé avec des systèmes de recommandation : une enquête sur les applications actuelles et l'impact potentiel, 2023)

1.3.2 Les défis et limitations de l'IA :

1. **Transparence et explication** : les algorithmes d'IA sont parfois considérés comme des "boîtes noires", ce qui rend difficile la compréhension de leurs recommandations.
2. **Données biaisées** : les modèles d'IA peuvent refléter les biais présents dans les données utilisées pour leur entraînement.
3. **Responsabilité médicale** : la question de la responsabilité en cas d'erreur diagnostique reste

Ces technologies ont prouvé leur rôle efficace dans l'interprétation des résultats et la résolution des problèmes liés aux ensembles de données. L'hôpital Apollo a également discuté du rôle de l'intelligence artificielle dans les soins aux patients atteints de cancer, en soulignant que ces technologies permettent de déterminer avec une plus grande précision le risque pour les individus de développer certains types de cancer, grâce à l'intégration de diverses sources de données. Elles contribuent également à identifier les personnes à haut risque, leur permettant de bénéficier de dépistages ciblés et de détecter la maladie à un stade précoce.

Dans une autre recherche, il a été mentionné que l'intelligence artificielle peut identifier de nouveaux biomarqueurs à partir des données, facilitant ainsi le dépistage des tumeurs, le diagnostic, la prédiction des traitements et les pronostics, offrant ainsi une meilleure prise en charge personnalisée des patients et améliorant leurs résultats cliniques.

Parmi ces technologies, on peut citer :

- IBM Watson Health : WFO peut être un outil utile pour la prise de décisions thérapeutiques contre le cancer du sein, en particulier dans les centres où les ressources spécialisées en oncologie du sein sont limitées. (Somashekhar, 2016) (Symposium, La plateforme IBM Watson for Oncology montre un haut degré de concordance avec les recommandations des médecins, 2016)
 - Tempus rend la médecine de précision une réalité grâce à la puissance et aux promesses des données et de l'intelligence artificielle. (Tempus, 2024)
 - Zebra Medical Vision : fondée en 2014, elle vise à améliorer le diagnostic médical grâce aux technologies d'apprentissage profond et à l'analyse des données médicales. (Vision)
- Google Health (DeepMind), une entreprise spécialisée dans l'intelligence artificielle appartenant à Google, a commencé à développer des applications pour les soins de santé, telles que l'application Streams, qui aide les médecins à surveiller les patients atteints d'insuffisance rénale aiguë. (DeepMind, 2019)

D'autres entreprises technologiques émergentes exploitent également l'intelligence artificielle pour améliorer le diagnostic médical, notamment dans le domaine de l'imagerie et de l'analyse automatisée. Parmi celles-ci, on peut citer :

- Aidoc, Arterys, DeepHealth, ...
- Aidoc, Arterys, DeepHealth, BenevolentAI et CureMetrix...

2 Problématique générale :

Le cancer du rein est une maladie dite « silencieuse », dont le diagnostic est souvent posé à un stade avancé en raison de l'absence de symptômes cliniques spécifiques dans les premières phases. Cette réalité rend son dépistage précoce difficile et complique l'efficacité de l'intervention thérapeutique. En Algérie, cette problématique est aggravée par une pénurie de compétences médicales spécialisées, ce qui conduit fréquemment à des erreurs de diagnostic, où l'on peut confondre une tumeur maligne avec une simple inflammation, entraînant parfois l'ablation inutile d'un rein sain.

Par ailleurs, l'analyse des données cliniques complexes – telles que les images médicales, les analyses biologiques et les données médicales – constituent un défi majeur en l'absence d'outils performants capables de corrélérer ces informations et d'en extraire des indicateurs pertinents. Cette situation prolonge les délais de diagnostic et de prise en charge, augmentant ainsi le risque de détérioration de l'état de santé du patient, voire de décès, en raison de retards dans la prise de décision médicale.

Dans ce contexte, la problématique suivante se pose :

Dans quelle mesure l'intelligence artificielle peut-elle contribuer à améliorer la précision et la rapidité du diagnostic du cancer du rein, en facilitant l'analyse de données cliniques hétérogènes, tout en imposant les délais de traitement et en améliorant les chances de survie des patients ?

3 Les Hypothèses :

- **Hypothèse 1 :**

Nous supposons que l'application des technologies d'intelligence artificielle pour analyser les informations cliniques des patients atteints d'un cancer du rein contribuerait à augmenter la précision du diagnostic médical et réduire le niveau d'erreur d'estimation de la pathologie.

- **Hypothèse 2 :**

Nous supposons que l'intégration de l'intelligence artificielle dans le processus de diagnostic et de traitement permettrait de réduire le temps nécessaire à la prise de décision médicale et thérapeutique, augmentant ainsi les chances d'une intervention précoce et, par conséquent, de sauver des vies.

Chapitre 01

Chapitre 01

Au cours des dernières années, nous avons observé un développement notable dans le domaine de la médecine, grâce à l'utilisation de l'intelligence artificielle dans différentes spécialités médicales. Cela se fait par le renforcement de la précision des diagnostics, l'amélioration des stratégies de traitement et l'optimisation des soins aux patients grâce aux interventions personnalisées et à la surveillance à distance. La revue met en lumière les tendances émergentes telles que l'intelligence artificielle explicable et l'apprentissage fédéré, tout en abordant des défis tels que le biais algorithmique, la transparence et l'impact sur la main-d'œuvre dans le domaine de la santé. L'établissement de lignes directrices réglementaires et la promotion de la collaboration entre médecins et développeurs d'intelligence artificielle sont essentiels pour surmonter ces obstacles. L'avenir de l'intelligence artificielle en médecine promet d'importantes améliorations des résultats pour les patients et de la prestation des soins de santé. (Basubrin., 2025)

Bien que l'intelligence artificielle puisse simplifier les tâches et réduire la charge de travail des médecins, elle ne remplace pas les praticiens humains. Les considérations éthiques et l'intégration de l'intelligence artificielle dans les pratiques de soins de santé sont essentielles pour maximiser ses avantages tout en faisant face aux défis et aux risques potentiels. (Srivastava, 2023) Il vise à intégrer à la fois les branches de la santé virtuelle et physique pour combler les lacunes en matière d'efficacité, de précision et de soins globaux aux patients, ce qui conduit à une transformation significative des pratiques médicales. (Almaeeni, 2024)

Bien que le cancer soit l'un des plus grands défis de la santé mondiale, l'utilisation de l'intelligence artificielle a entraîné des avancées notables et des résultats concrets. Cela se manifeste par un diagnostic précoce, qui est essentiel pour des interventions réussies. Les algorithmes d'intelligence artificielle détectent des anomalies fines que l'œil humain pourrait ignorer, en particulier dans les cancers du sein, des poumons et du foie (Fan Sun, 2025), ce qui permet un traitement en temps voulu. De plus, l'intelligence artificielle accélère la découverte de médicaments en analysant de grands ensembles de données pour identifier des candidats potentiels, réduisant ainsi le temps de développement et les coûts. Elle améliore également l'efficacité des essais cliniques en optimisant le recrutement des patients et la conception des essais. (Nizamullah FNU, 2024)

Parmi les principales techniques utilisées, on trouve le séquençage de nouvelle génération, l'imagerie médicale, l'apprentissage automatique et les réseaux de neurones, en

particulier des algorithmes tels que SVM, Naïve Bayes et CNN. (Amin Zadeh Shirazi, 6. The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide, 2024) (Mishra A. , 2023) (K. Aditya Shastry, 2021)

La revue met en lumière que l'apprentissage automatique et les algorithmes d'apprentissage profond, en particulier les réseaux de neurones, ont montré des résultats prometteurs dans le diagnostic du cancer, atteignant une performance similaire ou supérieure à celle des experts humains à travers différents types de cancer, améliorant ainsi la précision du diagnostic et les résultats des patients. (Professional, 2023) L'apprentissage profond, qui est une sous-catégorie de l'apprentissage automatique, améliore le diagnostic du cancer en analysant l'imagerie médicale (IRM, CT, radiographies) avec une précision et une rapidité accrues. (Analyse d'image : l'intelligence artificielle excelle dans l'interprétation des images médicales, y compris les scans radiologiques et les coupes pathologiques.) (B. Sakthivel, 2024) (Amin Zadeh Shirazi, The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide, 2024)

1 Deep learning :

L'apprentissage profond, également appelé « deeper learning », est un sous-domaine de l'apprentissage automatique qui se concentre sur la création et l'utilisation de simulations de réseaux neuronaux inspirées de la structure et du fonctionnement du cortex cérébral humain. Il s'agit d'un sous-ensemble de l'intelligence artificielle (IA) qui a historiquement prospéré grâce à son incroyable capacité d'apprentissage et de prédiction basée sur les grandes quantités de données actuellement disponibles. (Neethu Narayanan, 2024) A révolutionné de nombreux domaines, notamment le traitement du langage naturel, la reconnaissance vocale et la vision par ordinateur. Utilisant de nombreuses couches de réseaux neuronaux, il s'agit d'une composante de l'apprentissage automatique axée sur la modélisation de concepts de haut niveau. Le domaine de l'intelligence artificielle a été profondément transformé par cette technologie de pointe. L'apprentissage profond utilise des réseaux neuronaux artificiels pour traiter et acquérir d'énormes volumes d'informations, en s'inspirant de la composition et du fonctionnement du cerveau humain. Cette technologie a permis des avancées significatives dans des domaines tels que la vision par ordinateur, le traitement du langage naturel et la reconnaissance vocale. L'apprentissage profond est une technique efficace pour traiter des problèmes complexes et a le potentiel de révolutionner de nombreux secteurs d'activité dans les années à venir. (Soumik Kumar Mohanta S. , p. 2024)

1.1 Méthodes et approches dans l'apprentissage profond (Nicolas G. Polson, 2018) :

Réduction des données de haute dimension : L'apprentissage profond est principalement décrit comme une méthode de réduction des données de haute dimension. Cela implique la création de prédictions dans des modèles d'entrées-sorties, ce qui est essentiel pour gérer des ensembles de données complexes.

Couches hiérarchiques de caractéristiques latentes : L'une des méthodologies fondamentales de l'apprentissage profond est l'utilisation de couches hiérarchiques. Ces couches permettent au modèle d'apprendre et d'extraire des caractéristiques à plusieurs niveaux d'abstraction, ce qui est crucial pour des tâches telles que la reconnaissance d'images et le traitement du langage naturel.

Modélisation prédictive : L'article affirme que l'apprentissage profond est prédictif et non inférentiel. Cela signifie que l'accent est mis sur la réalisation de prédictions précises basées sur les données d'entrée, plutôt que sur la compréhension des relations ou des causes sous-jacentes.

Méthodologie de boîte noire : L'apprentissage profond est souvent désigné comme une méthodologie de boîte noire. Cela fait référence au fait que, bien que les modèles puissent fournir des estimations de fonctions de haute dimension, les processus internes et les décisions de ces modèles peuvent être opaques et difficiles à interpréter.

Approches algorithmiques : L'article examine les différentes perspectives algorithmiques liées à l'apprentissage profond. Bien que les algorithmes spécifiques ne soient pas détaillés dans le résumé, la mention des perspectives algorithmiques indique que l'article discute probablement des différents algorithmes utilisés dans l'apprentissage profond, tels que les réseaux de neurones convolutifs (CNN) et les réseaux de neurones récurrents (RNN).

Applications dans divers domaines : L'article liste également les applications réussies de l'apprentissage profond dans divers domaines, y compris l'intelligence artificielle (IA), le traitement d'images, la robotique et l'automatisation. Cela indique que les méthodes discutées ne sont pas seulement théoriques, mais ont également des applications pratiques dans des scénarios du monde réel. (Nicolas G. Polson, 2018)

1.2 Importance dans l'intelligence artificielle :

Technologie perturbatrice : L'apprentissage profond a considérablement changé le paysage de l'intelligence artificielle, repoussant les limites de ce qui peut être automatisé et développé (Soumik Kumar Mohanta A. G., 2024).

Potentiel futur : Sa capacité à traiter des problèmes complexes et à apprendre à partir de grands ensembles de données positionne l'apprentissage profond comme un pilier pour les avancées futures en intelligence artificielle (Machine Learning and Deep Learning: Algorithms and Applications, p. 2024).

Bien que l'apprentissage profond montre un potentiel immense, il est essentiel de considérer ses limites, telles que la nécessité de grands ensembles de données et de ressources computationnelles, ce qui peut entraver son accessibilité et son application dans certains contextes.

2 Machine Learning

L'apprentissage automatique (ML) est une technique centrale dans l'intelligence artificielle (IA), ayant un impact sur diverses industries et la vie quotidienne grâce aux avancées continues dans les méthodologies, ce qui la rend extrêmement influente à l'heure actuelle. (Pramod Gupta, 2024)

2.1 Concepts et techniques clés

Apprentissage supervisé : Cela implique des modèles entraînés sur des ensembles de données étiquetés pour prédire des résultats. Les techniques incluent les modèles linéaires, les arbres de décision et les réseaux de neurones (Sharma, 2024)

Apprentissage non supervisé : Il se concentre sur l'identification de motifs dans des données non étiquetées, telles que le regroupement et la réduction de dimensions (Géron, 2019)

Apprentissage par renforcement : Méthode par laquelle les agents apprennent à prendre des décisions en recevant des récompenses ou des pénalités en fonction de leurs actions. (Barto, 2018)

2.2 Applications du machine learning

Santé : Les algorithmes de machine learning aident au diagnostic et aux plans de traitement personnalisés, améliorant ainsi le processus de prise de décision pour les médecins (Ramírez., 2024)

Science des matériaux : La machine learning est utilisée pour prédire les propriétés des matériaux et améliorer les processus, ce qui réduit considérablement les coûts de calcul (M. Shiga, 2023)

Vie quotidienne : Des systèmes de recommandation à la reconnaissance vocale, la machine learning influence les activités quotidiennes et les processus industriels (Zhu, 2025).

Bien que la machine learning offre des possibilités transformantes dans divers secteurs, il fait également face à des défis tels que le surajustement et le besoin de méthodes d'évaluation robustes pour garantir la fiabilité du modèle (Khan, 2023) (K. Sudheer Babu, 2023)

Chapitre 02



L'intelligence artificielle (IA) est l'ensemble de techniques permettant à des systèmes informatiques de simuler des fonctions cognitives humaines telles que l'apprentissage, la résolution de problèmes ou la reconnaissance d'images aussi de langages et autre. Elle contient plusieurs sous-domaines, notamment le machine learning (l'apprentissage automatique), le deep learning (l'apprentissage profond), et les modèles de langage à grande échelle (LLM).

L'apprentissage profond (DP), en particulier, repose sur des réseaux de neurones artificiels multicouches capables d'apprendre des représentations complexes à partir de grandes quantités de données (LeCun, Bengio, & Hinton, 2015)[]. On site quelques architectures les plus utilisées :

- ✓ **Les autoencodeurs** : sont conçus pour apprendre une représentation compressée d'un input, on les utilise souvent pour la détection d'anomalies (normale / anormale).
- ✓ **Les réseaux convolutifs** : abrégé CNN, et on l'utilise généralement pour l'analyse d'images. Les CNN peuvent extraire automatiquement des caractéristiques visuelles telles que des formes, des textures et des structures.
- ✓ **Les réseaux récurrents** : abrégé RNN, c'est ce que j'utilise pour les données séquentielles, telles que le texte, les signaux biologiques, les séries chronologiques, etc.
- ✓ **Les Transformers** : basée sur le mécanisme d'attention, sont très performant pour le traitement du langage, et maintenant aussi pour les images.
- ✓ **Les GANs** : Generative Adversarial Networks, composés de deux réseaux s'affrontant – générateur et discriminateur, utilisés pour générer des images et des photos réalistes. (Frontiers in Digital Health, 2025)

Les frameworks les plus répandus incluent TensorFlow, Pytorch, Keras ou encore Scikit-learn. Leur accessibilité et leur performance ont facilité l'explosion des projets de recherche en IA dans de nombreux domaines, notamment la médecine.

1 Applications de l'IA en médecine :

1.1 Imagerie médicale : radiologie et histopathologie :

En raison de son efficacité dans l'analyse d'images médicales, IA d'images a révolutionné, l'IA a révolutionné le diagnostic médical .diagnostic médical. Utilisé en radiologie pour l'interprétation automatisée des radiographies, CT (tomodensitogrammes) ou IRM. Plusieurs des études ont démontré que certains modèles d'IA effectuent fonctionnent aussi bien, voire mieux,

que les spécialistes humains dans la détection de nodules pulmonaires ou de lésions mammaires au même niveau ou mieux que les spécialistes humains dans la détection des nodules pulmonaires ou des lésions mammaires. (McKinney, 2020).

Les réseaux de neurones convolutés (CNN) sont utilisés en histopathologie pour analyser numériquement les lames tissulaires afin d'identifier les anomalies cellulaires .utilisé en histopathologie pour analyser numériquement les lames tissulaires afin d'identifier les anomalies cellulaires. Particulier, Cela concerne notamment le diagnostic des cancers de la prostate, du sein et du rein. Cela concerne le diagnostic du cancer de la prostate, du sein et du rein. L'efficacité d'un réseau neuronal conjonctival (CNN) dans la classification des sous-types de cancer pulmonaire en fonction des lames H&E a été démontrée par une étude de Coudra et al. (2018), avec une précision supérieure à 90 %. (Coudray, 2018).

Dans notre projet, l'IA est appliquée à des images radiologiques et histopathologiques du rein afin de distinguer les tissus normaux, tumoraux. Ce domaine connaît une croissance rapide grâce à l'accessibilité des bases de données annotées et à la montée en puissance du calcul GPU.

1.2 Études déjà fait sur la différenciation tumeur / non-tumeur :

De nombreux travaux ont été menés pour classifier les images radiologiques et histologiques en classes « tumeur » vs « non tumeur » à l'aide des modèles d'AI :

- **Kurani et al. (2021)** ont utilisé un ResNet50 fine-tune pour classifier les images histologiques de rein, obtenant une précision de 94,7 %. (Kurani, 2021)
- **Campanella et al. (2019)** ont appliqué un CNN profond sur des lames entières (WSI) de prostate, avec une AUC de 0.98. (Campanella, 2019)
- **Zhao et al. (2020)** ont eu recours à des GANs pour générer des images synthétiques améliorant la robustesse des classifieurs CNN sur des tissus cancéreux mammaires. (Zhao & al, 2020)
- **Lu et al. (2023)** ont proposé un système combiné CNN + Transformer pour la classification de cancer colorectal, montrant une meilleure précision qu'un CNN seul. (Lu & al., 2023)

Tableau 1 : Tableau récapitulatif des exemples de travaux récents en classification tumeur / non-tumeur par IA.

Étude	Type de cancer	Modèle utilisé	Résultats principaux
Kurani et al. 2021	Rein	ResNet50	Accuracy = 94,7 %
Campanella et al., 2019	Prostate	WSI + CNN profond	AUC = 0.98
Zhao et al. 2020	Sein	GAN + CNN	Amélioration de la robustesse
Lu et al. 2023	Colorectal	CNN + Transformer	Gain de précision de +3 %

1.3 Aide au triage et télémédecine :

Des applications comme Babylon ou Ada utilisent l'IA pour guider les patients dans une phase pré-diagnostique. En télémédecine, l'IA sert à évaluer les urgences, prioriser les cas critiques et améliorer l'accès aux soins, notamment en zones sous-médicalisées. (Baker, et al., 2020)

Tableau 2 : Applications principales de l'IA en médecine et outils associés. (Baker, et al., 2020)

Domaine	Application	Modèle / Outil
Radiologie	Détection de tumeurs	CNN, ResNet50
Histopathologie	Analyse cellulaire	CNN, Autoencoder, GAN
NLP médical	Lecture de dossiers	BERT, GPT, BioBERT
Télémédecine	Triage automatique	XGBoost, Random Forest

2 Les modèles de deep learning en médecine :

2.1 Réseaux de neurones convolutifs (CNN) :

Les standards modèles pour l'analyse d'images sont les CNN. Pour l'analyse d'images sont les CNN. Constitué de tables convolutives qui extraient automatiquement les caractéristiques visuelles pertinentes (bords, textures, motifs). Et de couches entièrement connectées, il permet

une classification finale appropriée. Par exemple, un CNN peut apprendre à différencier un tissu tumoral d'un tissu normal en extrayant des motifs cellulaires spécifiques liés à la malignité.

Parmi les variantes les plus connues, on trouve VGGNet, InceptionNet et DenseNet. Des informations, et des canapés entièrement connectés permettent une catégorisation finale. Par exemple, en extrayant des modèles cellulaires spécifiques liés au cancer, un CNN peut apprendre à faire la distinction entre un tissu tumoral et un tissu normal. (Swart, Obaido, Jordan, & Ilono, 2025)

2.2 ResNet50 :

ResNet50 est un CNN profond avec 50 couches qui introduit des blocs résiduels, une innovation majeure permettant l'entraînement de réseaux très profonds sans dégradation 50 couches qui introduisent des blocs résiduels. Ces blocs utilisent des connexions sautées pour favoriser la propagation du gradient. Ce dernier a été appliqué avec succès à :

- La détection de métastases dans les ganglions lymphatiques (Camelyon challenge).
- La catégorisation des affections des lésions dermatologiques.
- L'examen histopathologique du rein, comme dans notre projet. (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2016)

2.3 Autoencodeurs et GANs :

« *Les autoencodeurs apprennent à reconstruire une entrée après l'avoir compressée. Dans un contexte médical, un autoencodeur entraîné uniquement sur des images normales aura des difficultés à reconstruire des images pathologiques, ce qui permet de détecter les anomalies. Exemple : détection précoce de tumeurs cérébrales invisibles à l'œil nu.*

Les GANs créent des images synthétiques à partir du bruit aléatoire, utiles dans plusieurs cas :

- *augmenter la taille des bases de données (data augmentation),*
- *harmoniser les images issues de différents centres médicaux (normalisation inter-centres),*
- *générer des images médicales réalistes pour l'entraînement de modèles sans données sensibles. »* (Shin, et al., 2018)

2.4 Transformers multimodaux :

Les Transformers, initialement conçus pour le traitement du langage naturel, sont désormais adaptés à l'analyse d'images (Vision Transformers) et de données mixtes (texte + image +

données cliniques). Leur capacité à traiter des dépendances longues dans les données permet d'intégrer des antécédents médicaux, des images d'imagerie, et des résultats biologiques pour établir un diagnostic plus global. (He, et al., 2022)

Exemple : un modèle multimodal peut croiser une uro-IRM avec un compte rendu clinique pour prédire la progression d'une maladie rénale.

Tableau 3 : Modèles de deep learning utilisés en médecine et leurs domaines d'application

Modèle	Description	Application principale
ResNet50	CNN profond avec blocs résiduels	Classification d'images médicales
Autoencoder	Reconstruction d'images	Détection d'anomalies
GAN	Génération d'images	Augmentation de données
Transformer	Apprentissage séquentiel multimodal	Synthèse de dossiers, NLP

3 Le modèle d'IA choisi pour notre étude :

Pour effectuer cette étude qui est basé sur l'IA pour avoir classé et faire la différence entre un rein avec tumeur et autre sain on a choisi le ResNet50.

3.1 Pourquoi :

Ce choix n'était pas aléatoire, mais il était basé sur une étude qui compare entre plusieurs modèles d'IA réalisées par (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018). Cela est dû au manque de temps et au volume de données utilisées, donc le traitement prend du temps, nous nous sommes donc principalement appuyés sur cette étude. Ils ont comparé les modèles d'IA et leurs rendements ou pourcentage d'accourcie(Accuracy) tout en gardant la même base de données (datasse) pour le découvrir le modèle le plus appropriée pour cet étude. Voici un tableau qui montre leur résultat :

Model	Patch-Wise		Image-Wise	
	Validation Set (%)	Test Set (%)	Validation Set (%)	Test Set (%)
ResNet50	90.68	94.03	89.58	97.50
Inception-V3	87.14	86.57	87.03	91.25

Tableau 4: Tableau récapitulatif des exemples de travaux récents en classification tumeur / non-tumeur par IA. (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018)

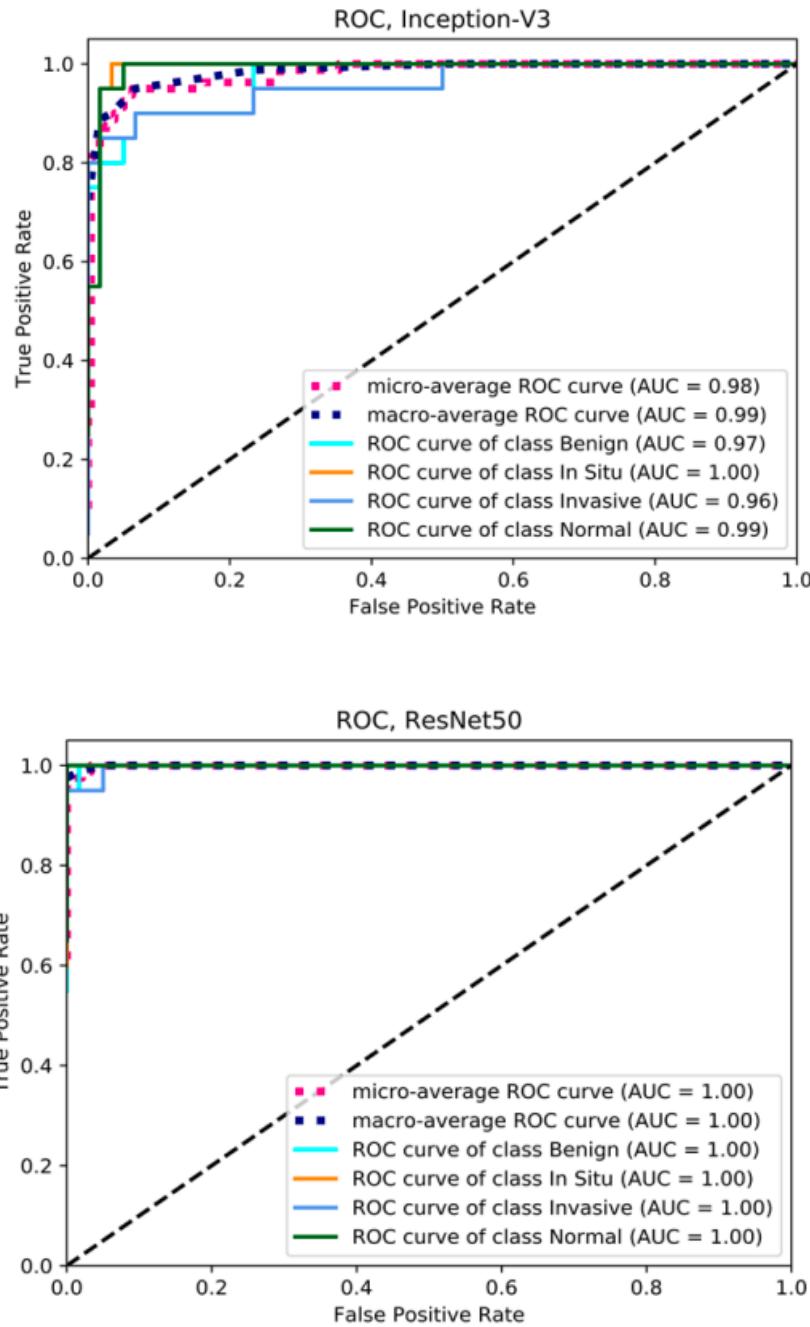


Figure 3: Courbes ROC pour l'ensemble de test non vu en utilisant les architectures ajustées de Google Inception-V3 et ResNet50. (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018)

Ils ont également calculé les courbes ROC (receiver operating characteristic) pour chaque réseau, représenté à la figure précédente. Les courbes ROC tracent le taux de vrais positifs (TPR) en fonction du taux de faux positifs (FPR) pour différents seuils de classification. Le TPR, également appelé sensibilité, représente la proportion d'échantillons correctement classés, tandis que le FPR, aussi appelé taux de fausses alertes, représente la proportion d'échantillons mal classés. Ainsi, la précision de classification a été mesurée par l'aire sous la courbe ROC (AUC),

une aire de 1 représentant une classification parfaite sur l'ensemble de test. Ils ont évalué les performances des réseaux pour chaque classe individuellement en calculant leurs courbes ROC et en déterminant les AUC correspondantes. La spécificité et la sensibilité globales de ResNet50 sont d'environ 99,9 %, tandis que celles d'Inception-V3 sont de 98 %. (Vesal, Ravikumar, Davari, Ellmann, & Maier, 2018). Ce qui confirme notre choix.

3.2 Structure du ResNet50 :

Le **ResNet50** (Residual Network à 50 couches) est une architecture phare du deep learning en vision par ordinateur. Proposé par He *et al.* (2015), il a remporté l'ImageNet ILSVRC 2015 en démontrant que les réseaux très profonds pouvaient être entraînés efficacement, grâce à l'introduction de blocs résiduels. (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2015)

3.2.1 Bloc résiduel :

Les blocs résiduels reposent sur les skip-connections : au lieu d'apprendre directement la fonction $H(x)$, le réseau apprend la fonction résiduelle $F(x) = H(x) - x$, puis reconstruit la sortie via $y = F(x) + x$. (contributors, 2024) ; Comme il est indiqué dans la figure ci-dessous.

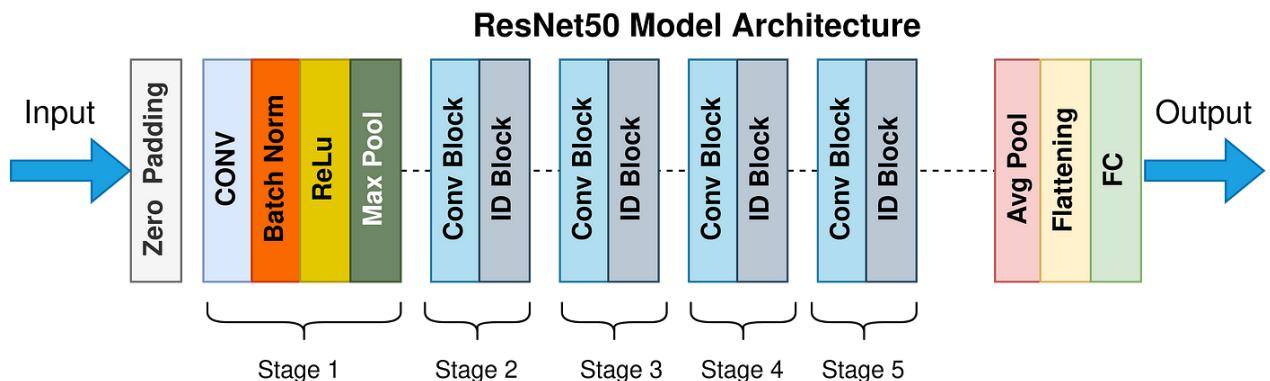


Figure 4 : architecture du modèle ResNet50. (Mishra D. , 2020)

- Couche 1×1 pour réduire la dimension (compression).
- Couche 3×3 .
- Couche 1×1 pour restaurer la dimension.
- Branche résiduelle (skip) ajoutée à la sortie. (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2015)

Cette architecture permet :

- d'éviter le vanishing gradient.

- de faciliter l'apprentissage dans les couches profondes.
- d'assurer une meilleure stabilité et précision sur des jeux de données complexes. (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2015)

3.2.2 Composition du ResNet50 :

ResNet50 se compose de :

- Une couche d'entrée Conv 7×7 + MaxPool
- 4 “stages” de blocs bottleneck :
 - Stage 1 : 3× bloc bottleneck
 - Stage 2 : 4× bloc
 - Stage 3 : 6× bloc
 - Stage 4 : 3× bloc
- Une couche Global Average Pooling
- Une tête fully connected (1000 classes ImageNet). (Agarwal, 2017)

Le total atteint 50 couches profondes, d'où le nom. Cette structure a démontré sa robustesse sur des tâches allant de la classification à la détection d'objets. (ResNet – Residual Neural Network Explained: Architecture, Variants, and Applications, s.d.)

Matériels et méthode

1 Données utilisées :

Dans le cadre de cette étude, on a exploité deux sources de données :

a) Des images radiologiques : librement accessibles sur la plateforme Kaggle, sou le lien <https://www.kaggle.com/datasets/jayaprakashpondy/kidney-cancer-image> destinées à la classification entre rein normale et tumoral. Le dataset initial a été téléchargé puis réorganisé manuellement en trois sous-ensembles :

- **Train (entraînement)** : 5887 images (79.99 %),
- **Validation** : 873 images (11.86 %),
- **Test** : 600 images (8.15 %).

Les données ont été équilibrées selon deux classes principales :

- **Normales** : 5077 images (68.98 %),
- **Tumeurs** : 2283 images (31.02 %).

Comme il est indiqué dans l'image ci-dessous :

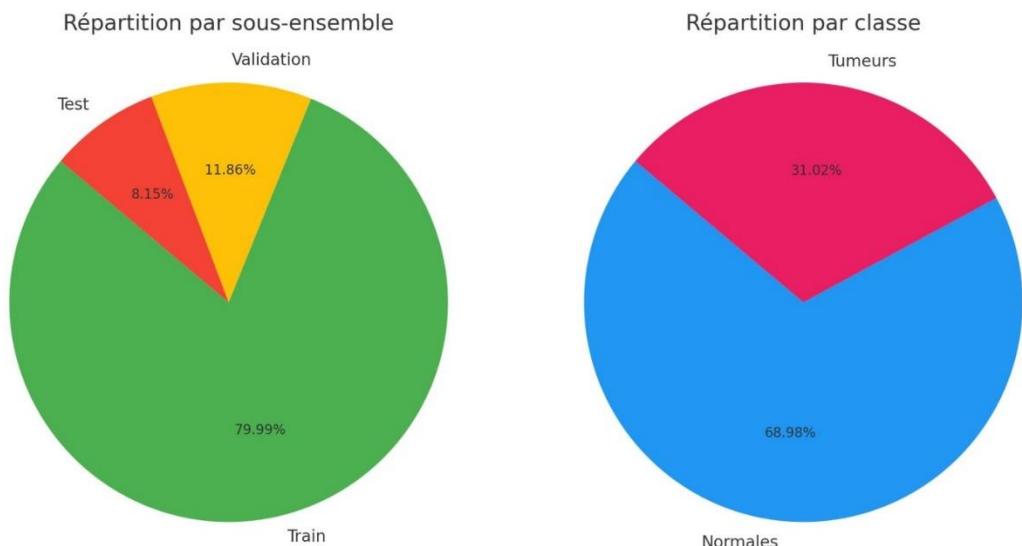


Figure 5 :Diagrammes circulaires illustrant la distribution des images radiologiques selon les sous-ensembles (Train, Validation, Test) et les classes (Normale, Tumeur).

b) Des images histopathologies : collectées directement auprès du service d'anatomopathologie de l'hôpital rénal de Daksi, 36 lames histologiques ont été analysées à l'aide d'un microscope optique « Leica» muni d'une caméra HD. Au total de **288 images** ont été enregistrées à **trois niveaux de grossissement différents** (4x, 10x et 40x). La majorité des cas concernait des **carcinomes à cellules claires** car il est le type histologique le plus fréquent

selon la classification de l'OMS, il représente environ 75 à 85% de tous les cancers des reins chez l'adulte (who classification of tumours editorial Board, 2022), et confirmé par le médecin du service.

Avec l'aide d'un médecin anatomopathologiste, les images ont été classées en deux catégories :

- **Normales** : tissus sans atypies ni signes d'infiltration tumorale.
- **Tumeurs** : zones contenant des cellules cancéreuses (souvent à cytoplasme clair).

Ces images ont ensuite été organisées en un dossier structuré par classe et par grossissement, la figure ci-dessous montre cette classification.

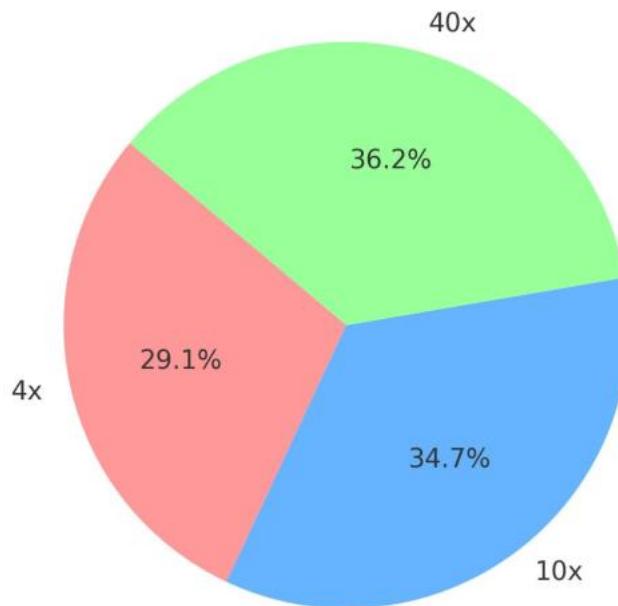


Figure 6 :Diagrammes circulaires illustrant la distribution des images histopathologies selon le grossissement (4x, 10x, 40x).

2 Prétraitement des données :

Les étapes de prétraitement diffèrent légèrement entre les deux types de données (radiologiques et histopathologies) en raison de leur nature intrinsèque, de leur qualité d'acquisition et de leur variabilité visuelle.

2.1 Radiologie :

Les images radiologiques téléchargées depuis la plateforme Kaggle sont généralement de qualité homogène et bien dimensionnées. On a appliqué les étapes suivantes :

- **Redimensionnement** des images à **224 x 224 pixels**, afin de standardiser la taille d'entrée du réseau ResNet50.

- **Conversion en RGB** (le cas échéant) pour garantir que toutes les images aient trois canaux (nécessaire pour les modèles pré-entraînés sur ImageNet).
- **Normalisation** des pixels : division des valeurs par 255 pour ramener les intensités dans l'intervalle [0, 1].
- **Encodage des étiquettes** (classe « normale » ou « tumeur ») via `to_categorical` pour obtenir des vecteurs one-hot de type [1, 0] ou [0, 1].

2.2 Histopathologie (microscope) :

Contrairement aux images issues de Kaggle, les données histopathologies présentent une plus grande variabilité liée au type de lames, à l'intensité des colorations (H&E), à la mise au point du microscope et aux conditions d'éclairage. Par conséquent, le traitement plus rigoureux a été nécessaire :

- **Redimensionnement** identique à 224 x 224 pixels, permettant une compatibilité directe avec ResNet50 (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2016).
- **Correction manuelle des erreurs d'image** (images floues, mal centrées, mal exposées).
- **Normalisation** des canaux RGB par division par 255. Cependant, certaines images nécessitaient un réajustement des contrastes et couleurs à l'aide de bibliothèques comme OpenCV ou PIL.
- **Équilibrage manuel** des classes : comme certaines classes (ex. normales à 40x) étaient sous-représentées, des images supplémentaires à 4x et 10x ont été incorporées pour renforcer l'échantillon.
- **Encodage** des étiquettes via `to_categorical` comme pour les données radiologiques.

La différence principale réside donc dans le **niveau de standardisation** initial des données. Les radiographies provenant de bases en ligne sont prétraitées et homogènes, tandis que les images histopathologies nécessitent une **normalisation plus fine** et un **contrôle qualité manuel** (Janowczyk & Madabhushi, 2016).

Dans les deux cas, les images ont été importées via la fonction `load_img` du module `keras.preprocessing.image`, converties avec `img_to_array` et ensuite transformées en tableaux NumPy normalisés.

3 Organisation et stratégie d'entraînement :

L'organisation de l'entraînement a été conçue de manière différenciée pour les données radiologiques et histopathologies, en tenant compte de leur spécificité structurelle, de leur diversité d'origine et de la stratégie de validation souhaitée.

3.1 Radiologie :

Pour les images radiologiques, une stratégie classique d'entraînement / validation / test a été adoptée, les données ayant été manuellement séparées par l'utilisateur en trois ensembles distincts :

- **Entraînement** (~70% des données),
- **Validation** (~15%),
- **Test** (~15%).

Ce découpage a été fait manuellement en amont, en tenant compte d'un bon équilibre entre les classes (tumeur et normale), ce qui a permis d'éviter l'usage de techniques plus complexes comme la validation croisée.

Les avantages de cette approche sont :

- Une répartition contrôlée des données,
- Un risque réduit de fuite de données entre les ensembles,
- Une meilleure adaptabilité à un ensemble relativement homogène et volumineux, comme c'est souvent le cas pour les bases radiologiques disponibles sur Kaggle (Litjens, et al., A survey on deep learning in medical image analysis, 2017).

L'entraînement a été effectué sur l'ensemble train, la performance évaluée sur val, puis le modèle final testé sur le lot test indépendant (données jamais vu).

3.2 Histopathologie :

L'approche adoptée repose sur un **entraînement indépendant** pour chaque niveau de grossissement (4x, 10x, 40x), dans le but d'évaluer l'impact de la résolution des images sur la performance du modèle de classification. Cette stratégie repose sur l'hypothèse que **chaque niveau de zoom capture des caractéristiques morphologiques distinctes** des tissus :

- À **4x**, on observe la structure globale et les architectures des lobules,
- À **10x**, l'observation se focalise sur les détails tissulaires intermédiaires,
- À **40x**, les détails cellulaires fins (noyaux, cytoplasme) sont mis en évidence (Abels, et al., 2019).

Pour chaque niveau, un schéma de validation croisée à 3 plis (3-fold K-Fold) a été mis en œuvre à l'aide de la classe KFold de sklearn.model_selection. Cette technique permet de :

- Réduire le biais lié au sous-échantillonnage,

- Garantir une évaluation robuste en utilisant l'ensemble des données pour l'entraînement et la validation à tour de rôle,
- Stabiliser les performances du modèle en moyenne.

3.3 Cas particulier du grossissement 40x :

Le niveau de grossissement 40x présentait un déséquilibre marqué entre les classes : très peu d'images normales étaient disponibles, ce qui risquait d'entraîner un surapprentissage du modèle sur la classe tumorale. Pour y remédier, une stratégie de compensation inter-niveaux a été adoptée :

- Les images normales issues des niveaux 4x et 10x ont été fusionnées au sein du dataset 40x,
- Ces ajouts ont permis d'enrichir la classe minoritaire (normale) tout en maintenant une certaine variabilité morphologique.
- Des vérifications ont été faites pour éviter les doublons et garantir la séparation entre les plis de validation et d'entraînement.

Cette solution pragmatique vise à rééquilibrer les classes tout en conservant la pertinence du niveau de détail visuel attendu à 40x. Elle s'aligne sur les pratiques recommandées dans la littérature concernant le traitement d'ensembles déséquilibrés en imagerie biomédicale (Janowczyk & Madabhushi, 2016) (Shamshad, et al., 2022).

4 Architecture du modèle ResNet50 appliqué :

Le modèle est fondé sur la célèbre architecture **ResNet50** (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2016), issue du domaine de la vision par ordinateur. Elle repose sur le **concept de connexions résiduelles** qui facilitent l'apprentissage dans les réseaux profonds, en contournant le problème du **gradient qui disparaît** dans les couches profondes. ResNet50 compte **50 couches convolutives**, ce qui lui permet d'extraire efficacement des **caractéristiques hiérarchiques** (des textures locales aux formes globales).

Dans notre travail, **deux modèles distincts** ont été entraînés :

- l'un à partir des **images radiologiques**.
- l'autre à partir des **images histopathologies**.

Les deux partagent la **même architecture de base**, mais leur **niveau de détail** et la **nature des signaux visuels** ont influencé les choix de certaines couches personnalisées.

4.1 Utilisation de la base ResNet50 :

Dans les deux cas, la base du réseau est importée via `keras.applications.ResNet50`, avec les **poids pré-entraînés sur ImageNet**, ce qui permet de **profiter du transfert d'apprentissage**. La couche de classification initiale a été **supprimée** (`include_top=False`) afin d'y ajouter une tête spécifique adaptée à notre tâche binaire (normal vs tumeur). La **base du modèle** (extracteur de caractéristiques) est restée gelée dans un premier temps (`trainable=False`) afin de stabiliser l'apprentissage des nouvelles couches et d'éviter de détruire les poids utiles déjà appris.

4.2 Tête personnalisée du modèle :

La tête ajoutée au-dessus de ResNet50 est commune aux deux modèles, mais chaque couche a été conçue avec une fonction spécifique :

```
Python
CopierModifier
x = GlobalAveragePooling2D () (base_model.output)
x = Dense (256, activation="relu", kernel_regularizer=l2 (0.001)) (x)
x = BatchNormalization () (x)
x = Dropout (0.5) (x)
x = Dense (128, activation="relu", kernel_regularizer=l2 (0.001)) (x)
x = BatchNormalization () (x)
x = Dropout (0.4) (x)
Output = Dense (2, activation="softmax") (x)
Model = Model (inputs=base_model.input, outputs=output)
```

4.3 Explication des couches ajoutées :

- `GlobalAveragePooling2D ()` : remplace le flattening classique. Elle résume chaque carte de caractéristiques par une moyenne, ce qui réduit les dimensions tout en conservant l'information essentielle. Elle est plus robuste au surapprentissage.
- `Dense (256) + activation="relu"` : couche entièrement connectée qui apprend des combinaisons complexes de caractéristiques extraites. La taille 256 est un compromis entre puissance de représentation et risque d'overfitting.
- `kernel_regularizer=l2 (0.001)` : régularisation L2 appliquée pour **limiter la complexité** du modèle et **réduire le surapprentissage**, surtout important dans le cas d'un petit jeu de données (notamment histopathologique).
- `BatchNormalization ()` : normalise l'activation des neurones à chaque batch pour **accélérer l'apprentissage** et stabiliser le réseau.
- `Dropout (0.5 puis 0.4)` : désactive aléatoirement un pourcentage de neurones à chaque itération pour **forcer le modèle à généraliser**. Le taux élevé de 0.5 est justifié par la **taille modeste du dataset** et la **complexité visuelle des images médicales**.

- Dense (128) : deuxième couche dense plus compacte qui affine les représentations apprises, avec les mêmes mécanismes de régularisation.
- Dense (2, activation="softmax") : couche finale de sortie avec deux neurones (tumeur vs normale), fournissant des **probabilités de classification binaire**.

5 Compilation, stratégie d'optimisation et fine-tuning :

Le modèle a été compilé avec l'**optimiseur Adam** ($lr = 0.001$), reconnu pour sa rapidité de convergence et sa robustesse dans les tâches de classification d'images médicales (Alzubaidi, et al., 2021). La fonction de perte choisie est la **categorical crossentropy**, adaptée aux tâches de classification binaire encodée en one-hot, et la métrique d'évaluation principale est l'**accuracy**, afin de mesurer directement la proportion d'images correctement classées.

Deux **callbacks dynamiques** ont été intégrés pour améliorer la stabilité de l'entraînement :

- ReduceLROnPlateau : cet outil ajuste automatiquement le taux d'apprentissage si la perte de validation ne s'améliore pas pendant trois epochs consécutives (patience = 3, factor = 0.5), ce qui permet d'explorer un espace de convergence plus fin à mesure que le modèle se stabilise.
- EarlyStopping : il interrompt l'apprentissage si aucune amélioration n'est observée sur les données de validation pendant sept epochs (patience = 7), afin de prévenir le surapprentissage et optimiser le temps de calcul.

Après cette phase initiale, une **stratégie de fine-tuning** a été appliquée conformément aux recommandations de (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2016) et (Shamshad, et al., 2022).

Les **20 dernières couches convolutionnelles** de la base ResNet50, initialement gelées, ont été **rendues entraînables** afin de permettre une spécialisation plus fine du modèle sur nos données médicales spécifiques. Durant cette phase, le taux d'apprentissage a été réduit à 1e-5 pour éviter des gradients trop brusques, préservant ainsi les poids pré-entraînés tout en autorisant une adaptation progressive aux caractéristiques des images histopathologiques.

6 Évaluation, visualisation et test final :

Pour chaque **pli de validation croisée**, les métriques suivantes ont été **calculées et tracées** :

- La **courbe de l'accuracy** pour l'entraînement et la validation, permettant de mesurer la capacité du modèle à classer correctement les images au fil des époques.
- La **courbe de la fonction de perte** pour suivre la convergence du modèle et détecter toute suradaptation.

Ces courbes ont été générées automatiquement via `matplotlib.pyplot` à partir des historiques d'entraînement retournés par la méthode `model.fit()`. Cela a permis de :

- Déterminer si le modèle convergeait correctement (baisse progressive de la perte),
- Identifier des cas de **surapprentissage (overfitting)** si la perte de validation augmentait malgré une baisse de la perte d'entraînement,
- Ajuster dynamiquement les paramètres via les callbacks `ReduceLROnPlateau` et `EarlyStopping`.

Un **ensemble de courbes a été sauvegardé pour chaque pli** et pour chaque niveau de grossissement (4x, 10x, 40x), ce qui a permis une comparaison directe des performances selon la résolution des images.

7 Bibliothèques et environnement technique :

7.1 Bibliothèques utilisées :

- ✓ TensorFlow 2.x et Keras
- ✓ NumPy
- ✓ Matplotlib
- ✓ scikit-learn
- ✓ Pillow (PIL)
- ✓ os, glob, pathlib
- ✓ tensorflow_addons
- ✓ pandas
- ✓ seaborn

7.2 Environnement d'exécution :

- ✓ **Langage** : Python 3.10,
- ✓ **Plateforme** : Windows 10,
- ✓ **IDE** : Jupyter Notebook,
- ✓ **Gestion d'environnement** : environnements virtuels configurés via conda

8 Interface web

Le développement de l'interface web du projet a été réalisé à l'aide du framework **Django** (nom du projet : *core*), en langage **Python**, avec l'intégration des technologies front-end **HTML**, **CSS**, **JavaScript**, et des bibliothèques modernes telles que **Chart.js**, **Bootstrap** et **Tailwind CSS** pour un rendu interactif, dynamique et responsive.

L'architecture du projet est composée de deux applications principales :

- **kidney_ai** : application centrale pour la gestion des modèles d'intelligence artificielle, le routage et l'affichage des résultats.
- **medical_data** : dédiée à la gestion et à l'analyse des données cliniques.

La page d'accueil propose une interface conviviale avec une section d'authentification. Après enregistrement, l'utilisateur peut accéder aux différentes sections via un menu supérieur interactif. Le fond d'écran alterne des visuels médicaux dynamiques, offrant une première impression esthétique et professionnelle.



Figure 7 : La page d'accueil de site web

L'interface est organisée autour de trois modules principaux :

1. **Radiologie** : permet à l'utilisateur de téléverser une image médicale (format .jpg), d'exécuter un modèle d'analyse basé sur l'IA, et d'afficher le résultat directement dans la même page.
2. **Histopathologie** : fonctionne de manière similaire à la section radiologique. L'utilisateur y importe des images histologiques afin d'évaluer automatiquement la présence éventuelle de cellules cancéreuses.

3. **Section Clinique & Statistiques** : Cette section présente une **analyse statistique scientifique** basée sur des données cliniques réelles. Elle affiche directement des **graphiques interactifs** (diagrammes en barres, en secteurs, etc.) permettant de visualiser les répartitions des patients, les tratement ou d'autres indicateurs pertinents liés au cancer du rein.

Un **système d'authentification personnalisé** a été mis en place pour les médecins : après connexion, le nom du médecin et sa photo apparaissent dans l'en-tête du site. Les résultats d'analyse sont également **stockés dans la base de données** pour un accès ultérieur.

Plusieurs bibliothèques ont été utilisées dans ce projet, notamment :

- **asgiref** (3.8.1) : pour la gestion asynchrone des requêtes.
- **django-widget-tweaks** (1.5.0) : pour améliorer le rendu des formulaires HTML.
- **pillow** (11.2.1) : pour le traitement des images médicales.
- **sqlparse** : pour l'analyse et la gestion des requêtes SQL.
- **tzdata** : pour la gestion des fuseaux horaires.

Les défurent Template professionnel :

- SoftLand
- Délicieux - Modèle Bootstrap de restaurant gratuit

Cette plateforme web constitue ainsi un pont fonctionnel entre les algorithmes d'intelligence artificielle et les professionnels de la santé, facilitant l'exploitation rapide des données médicales pour un diagnostic plus précis et une prise de décision optimisée.

Résultats et discussion

Results and discussion

1 Résultats radiologiques :

L'entraînement du modèle CNN basé sur **ResNet50** appliqué aux images radiologiques a été réalisé en deux phases : une première phase de **feature extraction**, où les couches convolutionnelles préentraînées sur ImageNet sont gelées, suivie d'un **fine-tuning** progressif. Cette stratégie s'appuie sur les principes du **transfert d'apprentissage**, bien connus pour améliorer les performances sur des bases de données médicales limitées (Shin, et al., 2016).

On a obtenu ces résultats :

- **Accuracy maximale en validation : 91,52 %**
- **Accuracy sur les données de test (non vues) : 88,83 %**
- **Loss final validation : 0.2966**

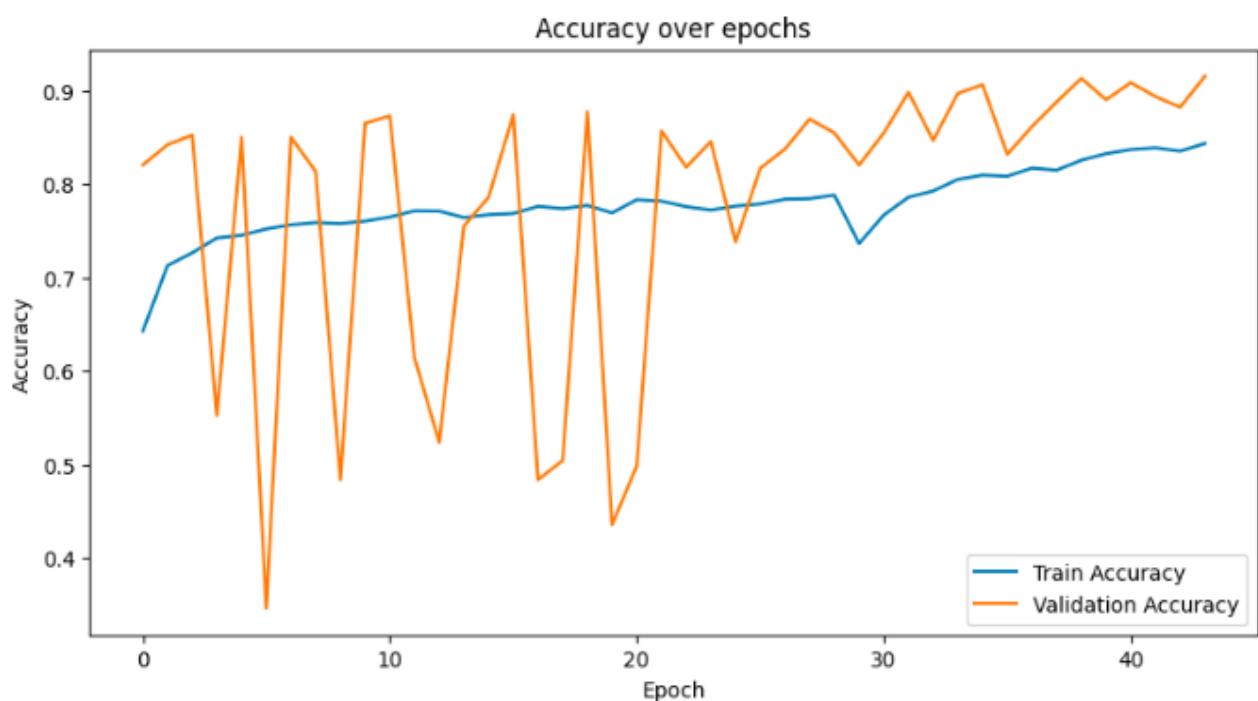


Figure 8 : Évolution de la précision d'entraînement et de validation au cours des époques pour le modèle ResNet50 avec fine-tuning sur un jeu de données médical binaire.

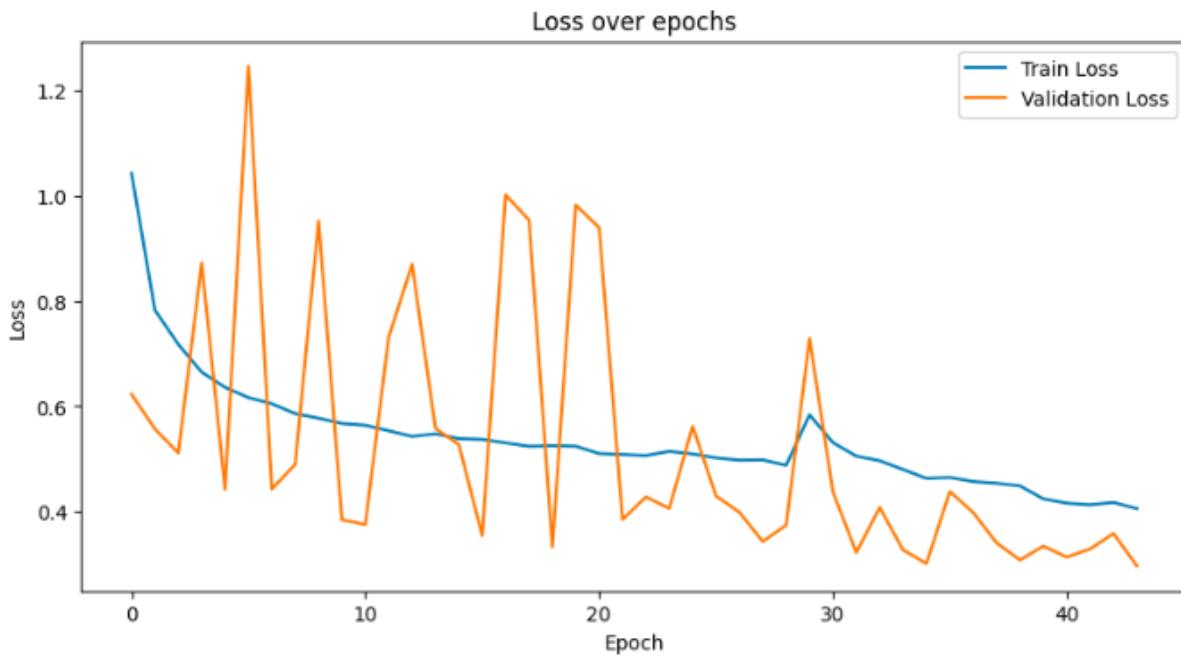


Figure 9: Évolution de la fonction de perte (loss) d’entraînement et de validation au cours des époques pour le modèle ResNet50 avec fine-tuning.

Le modèle a montré une amélioration progressive de ses performances, bien que marquée par quelques instabilités d’accuracy au fil des époques. Ces irrégularités sont fréquentes dans l’apprentissage profond sur des jeux médicaux complexes, souvent bruités ou déséquilibrés (Esteva, et al., 2017).

Le modèle ResNet50 préentraîné a permis une convergence rapide et précise, confirmant les résultats antérieurs sur d’autres modalités d’imagerie médicale (Raghu, Zhang, Kleinberg, & Bengio, 2019).

Les baisses temporaires d’accuracy observées à certaines époques pourraient être dues à la distribution inégale des classes ou à la présence d’artefacts visuels. Des approches comme l’augmentation de données ou le rééquilibrage peuvent corriger ces biais (Perez & Wang, 2017).

L’écart modéré entre l’accuracy de validation (91,52 %) et de test (88,83 %) indique une généralisation correcte du modèle, avec peu de signes d’**overfitting**.

Le temps d’entraînement moyen (environ 13 à 16 minutes par epoch) souligne la nécessité de ressources matérielles adéquates (GPU), surtout lors du fine-tuning. Cela reste cohérent avec la littérature pour des architectures profondes similaires (Litjens, et al., A survey on deep learning in medical image analysis, 2017).

❖ *Perspectives :*

Nous aspirons à affiner encore les performances, il serait pertinent d'expérimenter des techniques de régularisation plus poussées (Dropout, batch normalization), et aussi d'intégrer des méthodes de **data augmentation avancée** et de **segmentation préalable**.

2 Résultats histopathologiques :

Les performances des modèles entraînés sur les images histopathologiques capturées aux trois niveaux de grossissement (4x, 10x, 40x) ont été évaluées à l'aide d'une validation croisée à 3 plis (3-fold). Les résultats d'accuracy de validation obtenus pour chaque fold sont résumés dans les tableaux ci-dessous.

Grossissement	Fold 1	Fold 2	Fold 3	Moyenne	Écart-type
4x	74.47%	59.57%	78.72%	70.92%	±8.21%
10x	88.89%	80.00%	84.21%	84.37%	±3.65%
40x	92.31%	84.00%	88.00%	88.10%	±4.16%

Tableau 5 : Accuracy de validation par pli pour chaque grossissement.

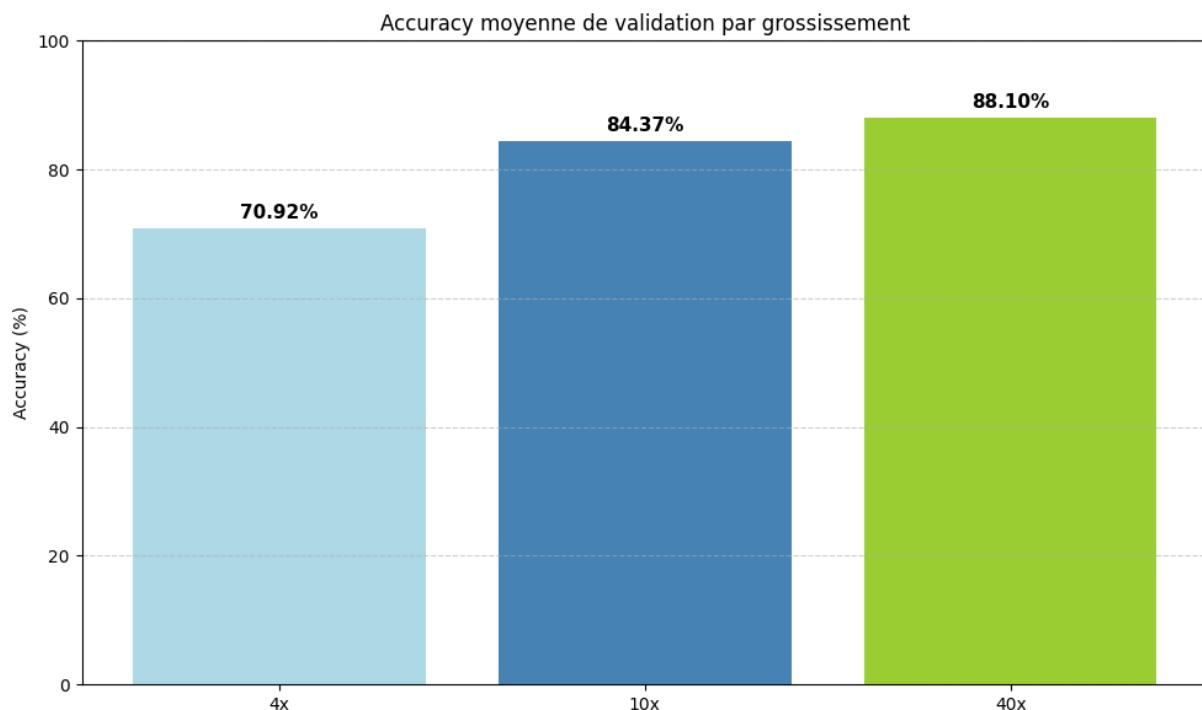


Figure 10: Courbe de validation accuracy par grossissement (résumé par moyenne des 3 folds).

Les résultats obtenus montrent une tendance nette en faveur des images à haute résolution : le grossissement 40x a obtenu une précision moyenne de validation de 88.10 %, supérieur au 10x (84.37 %) et nettement supérieur au 4x (70.92 %). Cette hiérarchie est cohérente avec le principe fondamental selon lequel une résolution plus élevée fournit plus de détails morphologiques aux modèles d'apprentissage profond ((Coudray, 2018); (Kurani, 2021)).

Le 40x offre une visibilité claire des structures cellulaires et sous-cellulaires (noyaux, cytoplasme, contours des cellules), ce qui permet à des modèles comme ResNet50 de détecter des motifs discriminants pour les tissus tumoraux. Le faible écart-type observé ($\pm 4.16 \%$) indique également une bonne stabilité du modèle à ce niveau.

En comparaison, les images 10x restent suffisamment riches pour permettre une détection précise des anomalies histologiques, expliquant leur performance moyenne de 84.37 %. Ce grossissement semble offrir un bon compromis entre précision, rapidité de traitement et charge mémoire, comme le suggèrent également d'autres études telles que celle de (Campanella, 2019).

À l'inverse, les performances du modèle entraîné avec les images 4x sont plus hétérogènes. La moyenne d'accuracy de 70.92 %, combinée à un écart-type élevé ($\pm 8.21 \%$), reflète à la fois une perte d'information visuelle due à la faible résolution, et des difficultés d'apprentissage du modèle. Ce phénomène est appuyé par la littérature qui souligne que les modèles CNN perdent en efficacité quand les patterns discriminants deviennent flous ou ambigus (Alzubaidi, et al., 2021).

Par ailleurs, un autre facteur explicatif du faible score du 4x est le déséquilibre partiel des classes : en absence d'images normales en 40x, une fusion artificielle des données normales 4x et 10x a été effectuée, ce qui a pu introduire une hétérogénéité dans le set d'entraînement du modèle 4x. La littérature confirme que les déséquilibres de classes et la qualité variable des données peuvent affecter négativement la capacité généralisable des modèles de deep learning (Shamshad, et al., 2022).

Enfin, l'approche de fine-tuning appliquée après un entraînement initial a permis une amélioration marginale des performances, en particulier pour le 40x.

Ce type de stratégie est reconnu pour améliorer la sensibilité des couches profondes aux motifs spécialisés (He, Zhang, Ren, & Sun, Deep Residual Learning for Image Recognition, 2016)

Conclusion

Conclusion

Ce travail de recherche s'inscrit dans les efforts visant à exploiter l'intelligence artificielle pour l'analyse des images médicales et des données cliniques, dans le but de soutenir le diagnostic du cancer du rein — l'un des cancers les plus complexes à classifier et à traiter. Nous avons opté pour une approche intégrée combinant trois perspectives complémentaires : la radiologie, l'histopathologie et les données cliniques, afin d'offrir une vision globale alliant imagerie et informations biologiques.

Au cours de ce travail, plusieurs défis réels ont été rencontrés, ayant eu un impact considérable sur le déroulement du projet et ses résultats. Ces obstacles peuvent être résumés comme suit :

- **Manque important de ressources et de données** : L'accès à des images médicales authentiques dans les bases de données internationales s'est avéré difficile, notamment en histopathologie, où les sources disponibles sont souvent limitées ou inadéquates pour le cancer du rein. Pour les images radiologiques, il nous a été impossible d'obtenir des clichés directement depuis les hôpitaux en raison de contraintes réglementaires et de confidentialité.
- **Contraintes temporelles** : Le temps imparti n'a pas permis de collecter un volume suffisant d'images, en particulier pour l'histopathologie, ce qui a limité la capacité du modèle à apprendre de manière approfondie. Nous avons tenté de pallier cela par l'utilisation d'images réelles disponibles, mais la taille de l'échantillon est restée restreinte.
- **Complexité du cancer du rein** : Ce type de cancer est particulièrement vaste et peu exploré, regroupant plus de 28 sous-types distincts aux caractéristiques histologiques et génétiques variées. Cette diversité a rendu la classification et la sélection des données particulièrement ardues. Nous avons donc concentré nos efforts sur un sous-type relativement fréquent.
- **Difficulté de classification des sous-types** : Le manque d'informations précises ou détaillées sur certains sous-types, notamment les plus rares, a compliqué les tâches de tri et d'analyse. De plus, certaines images disponibles n'étaient pas accompagnées d'une description anatomopathologique ou clinique claire.

Malgré ces limitations, les résultats obtenus par nos modèles d'IA se sont révélés prometteurs :

- Pour le modèle de radiologie, une précision de **91,52 %** a été atteinte lors de la phase d'entraînement, et environ **88 %** lors des tests sur des images inédites, démontrant une bonne capacité de généralisation.
- Concernant le modèle histopathologique, malgré un nombre restreint d'images, une précision supérieure à **62 %** a été obtenue, ce qui témoigne d'un potentiel certain, susceptible d'être amélioré par l'enrichissement de la base de données et l'amélioration de la qualité des images.

De manière générale, ce projet a démontré que l'intelligence artificielle peut efficacement contribuer au diagnostic du cancer du rein à travers l'analyse combinée des images et des données cliniques. Toutefois, une telle approche nécessite une alimentation riche et diversifiée en données, ainsi qu'une collaboration étroite avec les établissements médicaux pour l'obtention d'images authentiques et fiables. En outre, la nature complexe et hétérogène de cette pathologie exige le développement de modèles sur mesure pour chaque sous-type, objectif réalisable dans le cadre de recherches futures disposant de ressources et de temps suffisants.

En somme, ce travail constitue une première étape vers l'intégration de l'intelligence artificielle dans l'assistance au diagnostic du cancer du rein. Malgré les limites rencontrées, notamment en matière de données et de temps, les résultats obtenus montrent la pertinence et la faisabilité de notre approche. Ce projet ouvre ainsi la voie à des perspectives de recherche plus approfondies, notamment en enrichissant les bases de données, en collaborant étroitement avec les structures médicales et en adaptant les modèles aux spécificités de chaque sous-type tumoral. Ces efforts futurs permettront sans doute d'améliorer significativement la précision des diagnostics assistés par l'IA dans le domaine de l'oncologie rénale.

Bibliographies

- Machine Learning and Deep Learning: Algorithms and Applications. (2023). *Book Chapter*.
- 1, N. V. (2023). Obésité et carcinome rénal : mécanismes biologiques et perspectives. *Séminaires en biologie du cancer*, 21-33.
- A., L. (2023, 07 13). *Intelligence artificielle : va-t-elle remplacer le diagnostic médical ?* Retrieved from <https://www.inserm.fr/>.
- Abels, E., Pantanowitz, L., Aeffner, F., Zarella, M. D., van der Laak, J., Bui, M. M., . . . Kozlowski, C. (2019). Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *Journal of Pathology*, 249(3), pp. 286–294.
- Agarwal, D. (2017). *Understand and implement ResNets – CNN architectures for image classification*. Retrieved from CV-Tricks: <https://cv-tricks.com/keras/understand-implement-resnets>
- Almaeeni, H. (2024). Review on Artificial Intelligence in Medicine and Health. <https://ieeexplore.ieee.org/document/1080>, 59.
- Alzubaidi, L., Hasan, R. I., Yusuf, S. M., Oleiwi, S. R., Duan, Y., & Santamaría, J. (2021). Review of deep learning concepts, CNN architectures, challenges, applications, and future directions. *Journal of Big Data*, 8, p. 53.
- Amin Zadeh Shirazi, M. T. (2024). The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.
- Amin Zadeh Shirazi, M. T. (2024). 6. The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.
- Anne Beiter Arreskov, M. Å. (2019). L'impact du cancer sur les résultats du diabète. *Les fondamentaux de l'accès libre et de la recherche ouverte*.
- B. Sakthivel, P. T. (2024). Generative AI Models and Capabilities in Cancer Medical Imaging and Applications.
- Baker, A., Perov, Y., Middleton, K., Baxter, J., Mullarkey, D., Sangar, D., & Johri, S. (2020). A comparison of artificial intelligence and human doctors for the purpose of triage and diagnosis. *Frontiers in Artificial Intelligence*, 3, 543405. Retrieved from <https://doi.org/10.3389/frai.2020.543405>
- Barto, R. S. (2018). *Reinforcement Learning: An Introduction*. États-Unis (Cambridge, Massachusetts): MIT Press, collection Bradford Books.
- Basubrin., O. (2025). Current Status and Future of Artificial Intelligence in Medicine. *Cureus, Inc.*
- Campanella, G. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8), 1301–1309.
- Carlo, H. T. (2022). Cancer héréditaire et génétique du carcinome rénal. *Oncologie urologique*, 285–298.
- Chenais, G. (2023, juillet). Intelligence artificielle : va-t-elle remplacer le diagnostic médical ? fr, Inserm – La science pour la santé. Récupéré sur Intelligence artificielle : va-t-elle remplacer le

diagnostic médical ?: <https://www.inserm.fr/actualite/intelligence-artificielle-va-t-elle-replacer-le-diagnostic-medical/>

Conju-Probe, S. D.-U. (2022). *Le cancer du rein.*

contributors, W. (2024). *Residual neural network*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Residual_neural_network

Coudray, N. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24(10), 1559–1567.

DeepMind. (2019, July 31 31). *Blog post*. Récupéré sur Using AI to give doctors a 48-hour head start on life-threatening illness: <https://deepmind.google/discover/blog/using-ai-to-give-doctors-a-48-hour-head-start-on-life-threatening-illness/>

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), pp. 115–118.

Fan Sun, L. Z. (2025). Application progress of artificial intelligence in tumor diagnosis and treatment. *Frontiers in artificial intelligence*.

Frontiers in Digital Health. (2025). *Artificial intelligence in clinical medicine: a state-of-the-art overview* (1er ed.). Frontiers in Digital Health.

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. États-Unis (Sebastopol, Californie): O'Reilly Media, Inc.

He, K., Gan, C., Li, Z., Rekik, I., Yin, Z., Ji, W., . . . Shen, D. (2022). Transformers in Medical Image Analysis: A Review.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. Retrieved from arXiv: <https://arxiv.org/abs/1512.03385>

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. (IEEE, Ed.) Las Vegas, NV, USA, USA.

Imad-Addin, A. (2024). The Power of Artificial Intelligence for Improved Patient Outcomes, Ethical Practices and Overcoming Challenges. *10.61927/IGMIN222*, 597-600.

Intelligence artificielle : va-t-elle remplacer le diagnostic médical ? (2023, 07 13). Récupéré sur <https://www.inserm.fr/>.

Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7, p. 29.

JFR, T. L. (2024). Intelligence artificielle et diagnostic. *médical, quels enjeux éthiques...* *Journal de Radiologie – JFR*.

K. Aditya Shastry, H. A. (2021). Cancer diagnosis using artificial intelligence: a review. *Artificial Intelligence Review*.

K. Sudheer Babu, K. (2023). Applications de l'apprentissage automatique. *10.2174/9789815124514123010004*, 19-44.

- Khan, M. Z. (2023). Real world applications and research directions for machine learning: challenges and defies. *10.53555/ecb/2023.12.3.226*.
- Kurani, A. (2021). Deep learning for kidney cancer classification using histopathological images. *Computers in Biology and Medicine*, 136, 104706.
- Laurence Albiges 1, H. G. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. 881-891.
- Laurence Albiges 1, H. G. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. 881-891.
- Laurence Albiges, P. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. *The Lancet Oncology*.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, pp. 60-88.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., . . . Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, pp. 60-88.
- Lu, Y., & al. (2023). Hybrid deep learning model for colorectal cancer classification using CNN and Transformer. *IEEE Transactions on Medical Imaging*, 42(2), 1234-1245.
- Luis Pinto-Coelho1, 2. (2023). Comment l'intelligence artificielle façonne la technologie d'imagerie médicale : un aperçu des innovations et des applications. *Bioengineering*.
- M. Shiga, M. (2023). Machine Learning. 289-310.
- Mathilde Sibony 1, A. V. (2008). Carcinome rénal non à cellules claires. Mise à jour 2008 en pathologie tumorale rénale. *Annales de Pathologie*, 381-401.
- Mathilde Sibony 1, A. V. (2008). Carcinome rénal non à cellules claires. Mise à jour 2008 en pathologie tumorale rénale. *Annales de Pathologie*, 381-401.
- Matthew Young, M. a.-S. (2024). carcinome à cellules rénales. *Journal The Lancet*, 404 (10451), 476-491.
- Matthew Young, M. a.-S. (2024). carcinome à cellules rénales. *Journal The Lancet*, 404 (10451), 476-491.
- McKinney, S. M. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94.
- Mishra, A. (2023). 7. Advances Of AI And ML Technologies For Enhanced Cancer Prediction: An Exploratory Examination. *Journal of Namibian Studies : History Politics Culture*.
- Mishra, D. (2020). *The Annotated ResNet-50*. Retrieved from Medium: <https://medium.com/data-science/the-annotated-resnet-50-a6c536034758>

Naviguer dans le paysage des soins de santé avec des systèmes de recommandation : une enquête sur les applications actuelles et l'impact potentiel. (2023). *10.1109/ICCMC56507.2023.10083785*, 818-823.

Neethu Narayanan, N. (2024). Introduction à l'apprentissage profond.

Nicolas G. Polson, N. (2018). Apprentissage profond. *Accès libre*.

Nizamullah FNU, S. Z. (2024). AI in Healthcare: Breaking New Ground in the Management and Treatment of Cancer. *Asian journal of engineering, social and health*.

Patard, S. C.-J. (2007). *Le Cancer du rein* (Springer ed.). france: .

Patard, S. C.-J. (2007). *Le Cancer du rein* (9782287716508 ·2287716505 ed.). france.

Perez, L., & Wang, J. (2017, 12 13). *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*. Retrieved from arXiv: <https://doi.org/10.48550/arXiv.1712.04621>

Pramod Gupta, A. B. (2024). Machine Learning. *Book Chapter*.

Professional, P. S. (2023). 8. Artificial Intelligence in Cancer Diagnosis: A Systematic Review. *Physical Security Professional*.

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding Transfer Learning for Medical Imaging. In NeurIPS (Ed.), *Advances in Neural Information Processing Systems (NeurIPS 2019)*, Vol. 32. Vancouver, Canada.

Ramírez., D. J. (2024). Applications de l'apprentissage automatique dans le domaine de la santé : tendances actuelles et perspectives d'avenir. *Journal de la connaissance, de l'apprentissage et de la science et de la technologie (JKLST)*.

ResNet – Residual Neural Network Explained: Architecture, Variants, and Applications. (n.d.). Retrieved from Viso.ai: <https://viso.ai/deep-learning/resnet-residual-neural-network>

Sarah, B. (2024). Classification OMS 2022 des cancers du rein : focus sur le carcinome rénal papillaireClassification OMS 2022 des carcinomes à cellules rénales : Focus sur le carcinome papillaire à cellules rénales. *Annales de Pathologie*, 44(5), 314-322.

Shamshad, F., Khan, S., Zamir, S. W., Haris, M., Hayat, M., Shahbaz Khan, F., & Fu, H. (2022, 01 24). Transformers in Medical Imaging: A Survey.

Sharma, S. (2024). Apprentissage supervisé : une analyse approfondie. *Revue scientifique indienne de recherche*.

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., . . . Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), pp. 1285–1298.

Shin, H.-C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., . . . Michalski, M. (2018). Medical Image Synthesis for Data Augmentation and Anonymization using Generative Adversarial Networks.

Somashekhar. (2016, fév 01). *Watson for Oncology et recommandations de traitement du cancer du sein : accord avec un comité multidisciplinaire d'experts en tumeurs*. <https://www.ibm.com/>.

Soumik Kumar Mohanta, A. G. (2024). Deep Learning is a State-of-the-Art Approach to Artificial Intelligence. *Journal Article*.

Soumik Kumar Mohanta, S. (n.d.). L'apprentissage profond est une approche de pointe de l'intelligence artificielle.

Srivastava, R. (2023). Applications of Artificial Intelligence in Medicine. *Exploratory research and hypothesis in m...*.

Swart, T. G., Obaido, G., Jordan, M., & Ilono, P. (2025). Deep convolutional neural networks in medical image analysis: A review. *Information*, 16(3), 195.

Symposium. (2016, déc 14). *La plateforme IBM Watson for Oncology montre un haut degré de concordance avec les recommandations des médecins*. Retrieved from IBM : https://fr.newsroom.ibm.com/2016-12-14-La-plateforme-IBM-Watson-for-Oncology-montre-un-haut-degre-de-concordance-avec-les-recommandations-des-medecins?utm_source=chatgpt.com

Symposium. (2016, déc 14). *La plateforme IBM Watson for Oncology montre un haut degré de concordance avec les recommandations des médecins*. Récupéré sur IBM: https://fr.newsroom.ibm.com/2016-12-14-La-plateforme-IBM-Watson-for-Oncology-montre-un-haut-degre-de-concordance-avec-les-recommandations-des-medecins?utm_source=chatgpt.com

Tempus. (2024, juin). Tempus rend la médecine de précision une réalité grâce à la puissance et aux promesses des données et de l'intelligence artificielle. <https://www.youtube.com/watch?v=XLO8kPcMXJk>.

Tempus Roadshow Clips (June 2024) (2024). [Motion Picture].

Thomas Brüning 1, B. P. (2003). Risque de cancer des cellules rénales et exposition professionnelle au trichloréthylène : résultats d'une étude cas-témoins consécutive à Arnsberg, en Allemagne. *logo pubmed*, 274-85.

Tobacco and cancer. (1992). *European Journal of Cancer Prevention*, 211-214.

Venkatesh, N. (2023). Obésité et carcinome rénal : mécanismes biologiques et perspectives. *Séminaires en biologie du cancer*, 21-33.

Vesal, S., Ravikumar, ., Davari, ., Ellmann, ., & Maier, . (2018). Classification of breast cancer histology images using transfer learning.

Vision, Z. M. (s.d.). *About us*. Récupéré sur <https://zebramedical.com/about-us/>
who classification of tumours editorial Board. (2022). Urinary and Male Genital Tumours. (I. a. cancer, Ed.) Lyon, France.

Xiao Qiu Liu 1, G. P. (2019). Relations dose-réponse entre le tabagisme et le cancer du rein : revue systématique et méta-analyse. *PubMed - PMID: 31387065*, 86-93.

Yin, H. (2021). Consommation de viande rouge et transformée et conséquences sur le cancer : revue générale. *Chimie alimentaire*, 356.

Yuqing Zhang, M. P., Shufang Yang, P., & Zhen Wang, M. (2022). Association Between Diabetes and the Risk of Kidney Cancer: A Systematic Review and Meta-Analysis. *Clinical Diabetes*.

- Zhao, J., & al. (2020). Data augmentation using GANs for automated histopathology image analysis. *Medical Image Analysis*, 65, 101770.
- Zhu, C. Y. (2025). Applications de l'intelligence artificielle dans la vie quotidienne. 10.54254/2755-2721/2025.20074.
- Machine Learning and Deep Learning: Algorithms and Applications. (2023). *Book Chapter*.
- Almaeeni, H. (2024). Review on Artificial Intelligence in Medicine and Health. <https://ieeexplore.ieee.org/document/1080>, 59.
- Amin Zadeh Shirazi, M. T. (2024). The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.
- Amin Zadeh Shirazi, M. T. (2024). 6. The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.
- B. Sakthivel, P. T. (2024). Generative AI Models and Capabilities in Cancer Medical Imaging and Applications.
- Basubrin., O. (2025). Current Status and Future of Artificial Intelligence in Medicine. *Cureus, Inc*.
- Fan Sun, L. Z. (2025). Application progress of artificial intelligence in tumor diagnosis and treatment. *Frontiers in artificial intelligence*.
- K. Aditya Shastry, H. A. (2021). Cancer diagnosis using artificial intelligence: a review. *Artificial Intelligence Review*.
- Mishra, A. (2023). 7. Advances Of AI And ML Technologies For Enhanced Cancer Prediction: An Exploratory Examination. *Journal of Namibian Studies : History Politics Culture*.
- Neethu Narayanan, N. (2024). Introduction à l'apprentissage profond.
- Nicolas G. Polson, N. (2018). Apprentissage profond. *Accès libre*.
- Nizamullah FNU, S. Z. (2024). AI in Healthcare: Breaking New Ground in the Management and Treatment of Cancer. *Asian journal of engineering, social and health*.
- Pramod Gupta, A. B. (2024). Machine Learning. *Book Chapter*.
- Professional, P. S. (2023). 8. Artificial Intelligence in Cancer Diagnosis: A Systematic Review. *Physical Security Professional*.
- Soumik Kumar Mohanta, A. G. (2024). Deep Learning is a State-of-the-Art Approach to Artificial Intelligence. *Journal Article*.
- Soumik Kumar Mohanta, S. (s.d.). L'apprentissage profond est une approche de pointe de l'intelligence artificielle.

Srivastava, R. (2023). Applications of Artificial Intelligence in Medicine. *Exploratory research and hypothesis in m...*

Machine Learning and Deep Learning: Algorithms and Applications. (2023). *Book Chapter*.

1, N. V. (2023). Obésité et carcinome rénal : mécanismes biologiques et perspectives. *Séminaires en biologie du cancer*, 21-33.

A., L. (2023, 07 13). *Intelligence artificielle : va-t-elle remplacer le diagnostic médical ?* Retrieved from <https://www.inserm.fr/>.

Abels, E., Pantanowitz, L., Aeffner, F., Zarella, M. D., van der Laak, J., Bui, M. M., . . . Kozlowski, C. (2019). Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *Journal of Pathology*, 249(3), pp. 286–294.

Agarwal, D. (2017). *Understand and implement ResNets – CNN architectures for image classification.* Retrieved from CV-Tricks: <https://cv-tricks.com/keras/understand-implement-resnets>

Almaeeni, H. (2024). Review on Artificial Intelligence in Medicine and Health. <https://ieeexplore.ieee.org/document/1080>, 59.

Alzubaidi, L., Hasan, R. I., Yusuf, S. M., Oleiwi, S. R., Duan, Y., & Santamaría, J. (2021). Review of deep learning concepts, CNN architectures, challenges, applications, and future directions. *Journal of Big Data*, 8, p. 53.

Amin Zadeh Shirazi, M. T. (2024). The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.

Amin Zadeh Shirazi, M. T. (2024). 6. The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide. *Technology in Cancer Research & Treatment*.

Anne Beiter Arreskov, M. Å. (2019). L'impact du cancer sur les résultats du diabète. *Les fondamentaux de l'accès libre et de la recherche ouverte*.

B. Sakthivel, P. T. (2024). Generative AI Models and Capabilities in Cancer Medical Imaging and Applications.

Baker, A., Perov, Y., Middleton, K., Baxter, J., Mullarkey, D., Sangar, D., & Johri, S. (2020). A comparison of artificial intelligence and human doctors for the purpose of triage and diagnosis. *Frontiers in Artificial Intelligence*, 3, 543405. Retrieved from <https://doi.org/10.3389/frai.2020.543405>

Barto, R. S. (2018). *Reinforcement Learning: An Introduction*. États-Unis (Cambridge, Massachusetts): MIT Press, collection Bradford Books.

Basubrin., O. (2025). Current Status and Future of Artificial Intelligence in Medicine. *Cureus, Inc.*

Campanella, G. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8), 1301–1309.

Carlo, H. T. (2022). Cancer héréditaire et génétique du carcinome rénal. *Oncologie urologique*, 285–298.

Chenais, G. (2023, juillet). Intelligence artificielle : va-t-elle remplacer le diagnostic médical ? fr, Inserm – La science pour la santé. Récupéré sur Intelligence artificielle : va-t-elle remplacer le

diagnostic médical ?: <https://www.inserm.fr/actualite/intelligence-artificielle-va-t-elle-replacer-le-diagnostic-medical/>

Conju-Probe, S. D.-U. (2022). *Le cancer du rein.*

contributors, W. (2024). *Residual neural network*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Residual_neural_network

Coudray, N. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24(10), 1559–1567.

DeepMind. (2019, July 31 31). *Blog post*. Récupéré sur Using AI to give doctors a 48-hour head start on life-threatening illness: <https://deepmind.google/discover/blog/using-ai-to-give-doctors-a-48-hour-head-start-on-life-threatening-illness/>

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), pp. 115–118.

Fan Sun, L. Z. (2025). Application progress of artificial intelligence in tumor diagnosis and treatment. *Frontiers in artificial intelligence*.

Frontiers in Digital Health. (2025). *Artificial intelligence in clinical medicine: a state-of-the-art overview* (1er ed.). Frontiers in Digital Health.

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. États-Unis (Sebastopol, Californie): O'Reilly Media, Inc.

He, K., Gan, C., Li, Z., Rekik, I., Yin, Z., Ji, W., . . . Shen, D. (2022). Transformers in Medical Image Analysis: A Review.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. Retrieved from arXiv: <https://arxiv.org/abs/1512.03385>

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. (IEEE, Ed.) Las Vegas, NV, USA, USA.

Imad-Addin, A. (2024). The Power of Artificial Intelligence for Improved Patient Outcomes, Ethical Practices and Overcoming Challenges. *10.61927/IGMIN222*, 597-600.

Intelligence artificielle : va-t-elle remplacer le diagnostic médical ? (2023, 07 13). Récupéré sur <https://www.inserm.fr/>.

Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7, p. 29.

JFR, T. L. (2024). Intelligence artificielle et diagnostic. *médical, quels enjeux éthiques...* *Journal de Radiologie – JFR*.

K. Aditya Shastry, H. A. (2021). Cancer diagnosis using artificial intelligence: a review. *Artificial Intelligence Review*.

K. Sudheer Babu, K. (2023). Applications de l'apprentissage automatique. *10.2174/9789815124514123010004*, 19-44.

- Khan, M. Z. (2023). Real world applications and research directions for machine learning: challenges and defies. *10.53555/ecb/2023.12.3.226*.
- Kurani, A. (2021). Deep learning for kidney cancer classification using histopathological images. *Computers in Biology and Medicine*, 136, 104706.
- Laurence Albiges 1, H. G. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. 881-891.
- Laurence Albiges 1, H. G. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. 881-891.
- Laurence Albiges, P. (2023). Pembrolizumab plus lenvatinib comme traitement de première intention du carcinome rénal non à cellules claires avancé (KEYNOTE-B61) : un essai de phase 2 multicentrique à un seul bras. *The Lancet Oncology*.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, pp. 60-88.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., . . . Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, pp. 60-88.
- Lu, Y., & al. (2023). Hybrid deep learning model for colorectal cancer classification using CNN and Transformer. *IEEE Transactions on Medical Imaging*, 42(2), 1234-1245.
- Luis Pinto-Coelho1, 2. (2023). Comment l'intelligence artificielle façonne la technologie d'imagerie médicale : un aperçu des innovations et des applications. *Bioengineering*.
- M. Shiga, M. (2023). Machine Learning. 289-310.
- Mathilde Sibony 1, A. V. (2008). Carcinome rénal non à cellules claires. Mise à jour 2008 en pathologie tumorale rénale. *Annales de Pathologie*, 381-401.
- Mathilde Sibony 1, A. V. (2008). Carcinome rénal non à cellules claires. Mise à jour 2008 en pathologie tumorale rénale. *Annales de Pathologie*, 381-401.
- Matthew Young, M. a.-S. (2024). carcinome à cellules rénales. *Journal The Lancet*, 404 (10451), 476-491.
- Matthew Young, M. a.-S. (2024). carcinome à cellules rénales. *Journal The Lancet*, 404 (10451), 476-491.
- McKinney, S. M. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94.
- Mishra, A. (2023). 7. Advances Of AI And ML Technologies For Enhanced Cancer Prediction: An Exploratory Examination. *Journal of Namibian Studies : History Politics Culture*.
- Mishra, D. (2020). *The Annotated ResNet-50*. Retrieved from Medium: <https://medium.com/data-science/the-annotated-resnet-50-a6c536034758>

Naviguer dans le paysage des soins de santé avec des systèmes de recommandation : une enquête sur les applications actuelles et l'impact potentiel. (2023). *10.1109/ICCMC56507.2023.10083785*, 818-823.

Neethu Narayanan, N. (2024). Introduction à l'apprentissage profond.

Nicolas G. Polson, N. (2018). Apprentissage profond. *Accès libre*.

Nizamullah FNU, S. Z. (2024). AI in Healthcare: Breaking New Ground in the Management and Treatment of Cancer. *Asian journal of engineering, social and health*.

Patard, S. C.-J. (2007). *Le Cancer du rein* (Springer ed.). france: .

Patard, S. C.-J. (2007). *Le Cancer du rein* (9782287716508 ·2287716505 ed.). france.

Perez, L., & Wang, J. (2017, 12 13). *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*. Retrieved from arXiv: <https://doi.org/10.48550/arXiv.1712.04621>

Pramod Gupta, A. B. (2024). Machine Learning. *Book Chapter*.

Professional, P. S. (2023). 8. Artificial Intelligence in Cancer Diagnosis: A Systematic Review. *Physical Security Professional*.

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding Transfer Learning for Medical Imaging. In NeurIPS (Ed.), *Advances in Neural Information Processing Systems (NeurIPS 2019)*, Vol. 32. Vancouver, Canada.

Ramírez., D. J. (2024). Applications de l'apprentissage automatique dans le domaine de la santé : tendances actuelles et perspectives d'avenir. *Journal de la connaissance, de l'apprentissage et de la science et de la technologie (JKLST)*.

ResNet – Residual Neural Network Explained: Architecture, Variants, and Applications. (n.d.). Retrieved from Viso.ai: <https://viso.ai/deep-learning/resnet-residual-neural-network>

Sarah, B. (2024). Classification OMS 2022 des cancers du rein : focus sur le carcinome rénal papillaireClassification OMS 2022 des carcinomes à cellules rénales : Focus sur le carcinome papillaire à cellules rénales. *Annales de Pathologie*, 44(5), 314-322.

Shamshad, F., Khan, S., Zamir, S. W., Haris, M., Hayat, M., Shahbaz Khan, F., & Fu, H. (2022, 01 24). Transformers in Medical Imaging: A Survey.

Sharma, S. (2024). Apprentissage supervisé : une analyse approfondie. *Revue scientifique indienne de recherche*.

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., . . . Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), pp. 1285–1298.

Shin, H.-C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., . . . Michalski, M. (2018). Medical Image Synthesis for Data Augmentation and Anonymization using Generative Adversarial Networks.

Somashekhar. (2016, fév 01). *Watson for Oncology et recommandations de traitement du cancer du sein : accord avec un comité multidisciplinaire d'experts en tumeurs*. <https://www.ibm.com/>.

Soumik Kumar Mohanta, A. G. (2024). Deep Learning is a State-of-the-Art Approach to Artificial Intelligence. *Journal Article*.

Soumik Kumar Mohanta, S. (n.d.). L'apprentissage profond est une approche de pointe de l'intelligence artificielle.

Srivastava, R. (2023). Applications of Artificial Intelligence in Medicine. *Exploratory research and hypothesis in m...*.

Swart, T. G., Obaido, G., Jordan, M., & Ilono, P. (2025). Deep convolutional neural networks in medical image analysis: A review. *Information*, 16(3), 195.

Symposium. (2016, déc 14). *La plateforme IBM Watson for Oncology montre un haut degré de concordance avec les recommandations des médecins*. Retrieved from IBM : https://fr.newsroom.ibm.com/2016-12-14-La-plateforme-IBM-Watson-for-Oncology-montre-un-haut-degre-de-concordance-avec-les-recommandations-des-medecins?utm_source=chatgpt.com

Symposium. (2016, déc 14). *La plateforme IBM Watson for Oncology montre un haut degré de concordance avec les recommandations des médecins*. Récupéré sur IBM: https://fr.newsroom.ibm.com/2016-12-14-La-plateforme-IBM-Watson-for-Oncology-montre-un-haut-degre-de-concordance-avec-les-recommandations-des-medecins?utm_source=chatgpt.com

Tempus. (2024, juin). Tempus rend la médecine de précision une réalité grâce à la puissance et aux promesses des données et de l'intelligence artificielle. <https://www.youtube.com/watch?v=XLO8kPcMXJk>.

Tempus Roadshow Clips (June 2024) (2024). [Motion Picture].

Thomas Brüning 1, B. P. (2003). Risque de cancer des cellules rénales et exposition professionnelle au trichloréthylène : résultats d'une étude cas-témoins consécutive à Arnsberg, en Allemagne. *logo pubmed*, 274-85.

Tobacco and cancer. (1992). *European Journal of Cancer Prevention*, 211-214.

Venkatesh, N. (2023). Obésité et carcinome rénal : mécanismes biologiques et perspectives. *Séminaires en biologie du cancer*, 21-33.

Vesal, S., Ravikumar, ., Davari, ., Ellmann, ., & Maier, . (2018). Classification of breast cancer histology images using transfer learning.

Vision, Z. M. (s.d.). *About us*. Récupéré sur <https://zebramedical.com/about-us/>
who classification of tumours editorial Board. (2022). Urinary and Male Genital Tumours. (I. a. cancer, Ed.) Lyon, France.

Xiao Qiu Liu 1, G. P. (2019). Relations dose-réponse entre le tabagisme et le cancer du rein : revue systématique et méta-analyse. *PubMed - PMID: 31387065*, 86-93.

Yin, H. (2021). Consommation de viande rouge et transformée et conséquences sur le cancer : revue générale. *Chimie alimentaire*, 356.

Yuqing Zhang, M. P., Shufang Yang, P., & Zhen Wang, M. (2022). Association Between Diabetes and the Risk of Kidney Cancer: A Systematic Review and Meta-Analysis. *Clinical Diabetes*.

Zhao, J., & al. (2020). Data augmentation using GANs for automated histopathology image analysis. *Medical Image Analysis*, 65, 101770.

Zhu, C. Y. (2025). Applications de l'intelligence artificielle dans la vie quotidienne. [10.54254/2755-2721/2025.20074](https://doi.org/10.54254/2755-2721/2025.20074).

Analyse des données cliniques et histopathologiques par des approches basées sur l'intelligence artificielle pour le diagnostic du cancer de rein

Mémoire pour l'obtention du diplôme de Master en Bio-informatique

Résumé

Ce travail de recherche vise à développer un modèle d'intelligence artificielle contribuant à améliorer la rapidité du diagnostic histologique du cancer du rein, un type de cancer silencieux dont la détection précoce reste difficile. Le travail s'articule autour de l'analyse de données provenant de sources multiples, incluant des images médicales (radiologiques et histopathologiques) ainsi que des données cliniques réelles, en utilisant des techniques avancées d'apprentissage profond (Deep Learning) via le modèle ResNet50.

Les données ont été collectées à partir de la plateforme Kaggle et le service d'anatomopathologie de la clinique rénale de Daksi Constantine. Elles ont été soigneusement classées et prétraitées avant l'entraînement du modèle.

Une interface web interactive a également été développée pour permettre à l'utilisateur d'obtenir automatiquement les résultats d'analyses histologiques.

Les résultats que nous avons obtenus ont montré une grande efficacité du modèle dans la distinction entre les tissus sains et les tissus atteints, ce qui souligne la capacité de l'intelligence artificielle à soutenir la décision médicale et à réduire les délais diagnostic.

Ce projet constitue une avancée significative vers l'intégration des technologies d'intelligence artificielle dans le domaine de la santé en Algérie, en renforçant les chances de dépistage précoce et en améliorant la qualité des soins apportés aux patients.

Mots-clefs : l'intelligence artificielle, cancer de rein, Analyse des données cliniques et histopathologiques

Laboratoires de recherche : laboratoire de Centre de Recherche en Biotechnologie « CRBt »

Établissement de travail : Clinique Rénale El Daksi, Constantine

Président du jury : Dr. Djezzar Nedjma (PROF- UConstantine1 Frères Mentouri).

Encadrant : Dr. Bensaada Mostefa (PROF- UFM Constantine 1).

Co-Encadrant : Dr. Chaita Kamel (Médecin spécialisé en anatomopathologie)

Examinateur(s): Dr. Krid Adel (PROF - UFM Constantine 1).