

THEME :
DETECTION DU CANCER DU SEIN :
CONCEPTION D'UNE APPLICATION
D'INTELLIGENCE ARTIFICIELLE BASEE
SUR LE DEEP LEARNING

Rapport de Stage en vue de l'obtention du :

Bachelor en Intelligence Artificielle et Big Data

Rédigé et soutenu par :

Kamguem Mande Ange Muriel

Encadrant académique :

Mr Abdouraman Bouba Dalil

Encadrant professionnel :

Mr Peguy Njamo

Année Académique : 2022/2023

RESUME

Le cancer du sein a été documenté depuis l'Antiquité, mais les premières descriptions cliniques détaillées de tumeurs du sein remontent au temps de l'Égypte ancienne, vers 1600 avant J.-C. Cependant, la compréhension moderne du cancer du sein et son association avec des cellules cancéreuses n'ont vraiment émergé que plus tard, au 18^e et 19^e siècle. Des progrès significatifs dans la compréhension de la maladie et de son traitement ont été réalisés au cours du 20^e siècle.

Le cancer du sein est l'une des principales pathologies oncologiques affectant les femmes dans le monde. Avec le temps, le nombre de cas a continué d'augmenter. Selon les données de l'Organisation Mondiale de la Santé, les cas de cancer du sein au Cameroun ont augmenté avec au moins 39000 camerounais vivant actuellement avec le cancer du sein. Les adultes sont les plus touchés, avec 15.262 nouveaux cas chez les plus de 15 ans.

La croissance rapide de cette maladie a entraîné un besoin accru de méthodes de détection précoces. Étant donné les limites des ressources et la gravité du cancer du sein, nous avons proposé un modèle d'apprentissage profond pour aider les radiologues et les cliniciens à détecter les tumeurs mammaires à partir de mammographies.

Notre objectif dans cette étude était de mettre en œuvre une technique de Deep Learning pour la détection du cancer du sein. Nous avons opté pour les réseaux neuronaux convolutifs (CNN), reconnus pour leur capacité en classification d'images.

Les résultats démontrent que notre modèle peut détecter les cas de cancer du sein avec une précision de 71%.

Mots clés : Cancer du sein, Machine Learning, Apprentissage supervisé, Réseaux de neurones, Réseaux de neurones convolutif, Classification.

ABSTRACT

Breast cancer has been documented since ancient times, but the first detailed clinical descriptions of breast tumors date back to ancient Egypt, around 1600 BC. However, modern understanding of breast cancer and its association with cancer cells only emerged much later, in the 18th and 19th centuries. Significant progress in understanding the disease and its treatment has been made during the 20th century.

In Cameroon, the first cases of breast cancer were identified well before 2020. Breast cancer is one of the leading oncological pathologies affecting women worldwide. Over time, the number of cases has continued to increase. According to the World Health Organization's data, cases of breast cancer in Cameroon has increased, with at least 39,000 Cameroonians currently living with breast cancer. Adults are the most affected, with 15,262 new cases among those over 15 years old. .

The rapid growth of this disease has led to an increased need for early detection methods. Given the limitations of resources and the severity of breast cancer, we proposed a deep learning model to help radiologists and clinicians detect breast tumors from mammograms .

Our goal in this study was to implement a Deep Learning technique for breast cancer detection. We opted for Convolutional Neural Networks (CNN), recognized for their image classification capabilities .

The results demonstrate that our model can detect cases of breast cancer with an accuracy of **71%** .

Keywords: Breast Cancer, Machine Learning, Supervised Learning, Neural Networks, Convolutional Neural Networks, Classification.

SOMMAIRE

RESUME	2
ABSTRACT	3
SOMMAIRE	4
ACRONYMES ET SIGLES	5
LISTES DES FIGURES	7
INTRODUCTION GENERALE	7
I. Contexte général de l'étude	8
II. Problématique de l'étude	9
a. Présentation du problème	9
b. Problème général	9
c. Problèmes spécifiques	9
III. Hypothèse de l'étude	9
a. Hypothèse générale	9
b. Hypothèses spécifiques	9
IV. Objectif de l'étude	10
a. Objectif général l'étude	10
b. Objectif Spécifique de l'étude	10
c. Justification de l'Etude	10
d. Délimitation de l'étude	10
e. Plan du mémoire	11
CHAPITRE I : CADRE THEORIQUE ET CONCEPTUEL	
Introduction	13
I. Cadre théorique et état d'art	13
II. Historique	14
III. Cadre Règlementaire	16
Conclusion	17
CHAPITRE II : Méthodologie de recherche	
Introduction	19
I. Nature de recherche	19
II. Variable de recherche	19
a. Définition conceptuelle	19

Instruments de recherche	22
Conclusion	24
CHAPITRE III : PRESENTATION D'INONI TECH ET DES DONNEES COLLECTEES	
Introduction	26
I. Présentation de l'entreprise	26
a. Produits et services de l'entreprise	26
b. Localisation de l'entreprise	26
c. Mission	27
d. Organigramme de l'entreprise	28
II. Données collectées	28
Conclusion	29
CHAPITRE IV : ANALYSE DIAGNOSTIQUE DE LA SITUATION ET PROPOSITION D'INTERVENTION	
Introduction	31
I. Présentation et analyse de la situation	31
II. Intervention proposée et justification	32
III. Objectifs de l'intervention - projet envisagé	33
a. Objectif général	33
b. Objectifs spécifiques	33
IV. Composante de l'intervention envisagée	33
a. Stratégies d'action et contenu : Conception d'un modèle Deep Learning pour la détection du Cancer du sein	33
b. Résultats obtenus et discussion	35
V. Faisabilité	38
1. Technique	38
Conclusion	39
CONCLUSION GENERALE	40
REFERENCES WEBOGRAPHIQUES	41
REFERENCES BIBLIOGRAPHIQUES	41
TABLE DE MATIERES	42

ACRONYMES ET SIGLES

IA : Intelligence artificielle

ML: Machine Learning

DL: Deep Learning

OMS : Organisation mondial de la Sante

ReLU : Unités Rectifié linéaires

CNN ou Conv : Convolutional Neural Network, Réseau de neurones convoluti

LISTES DES FIGURES

Figure 1 : Algorithme de Machine Learning	20
Figure 2 : <i>Intelligence Artificielle et ses Composants</i>	21
Figure 3 : <i>Architecture d'un CNN</i>	22
Figure 4 : <i>Localisation de Inoni Tech</i>	27
Figure 5 : Organigramme de l'entreprise	28
Figure 6 : <i>Mammographie</i>	29
Figure 7 : Présentation de l'évolution du cancer du sein	31
Figure 8 : création et optimisation du modèle	34
Figure 9 : structure du modèle.....	35
Figure 10 : résultat du modèle.....	36
Figure 11 : les epochs.....	36
Figure 12 : code pour la matrice de confusion	37
Figure 13 : matrice de confusion.....	37
Figure 14 : fonction de prédiction	38

LISTES DES TABLEAUX

Tableau 1 : <i>Ressources logicielle</i>	38
Tableau 2: <i>Ressources matérielles</i>	39

INTRODUCTION GENERALE

A la fin du cursus Bachelor à l'Université Supérieure Française Keyce Informatique et Intelligence Artificielle, chaque étudiant se doit d'effectuer un stage académique en vue de l'obtention de son diplôme de Licence. Ce stage offre la possibilité à chaque étudiant de mettre en pratique les connaissances acquises au cours de sa formation, d'acquérir de nouvelles compétences et aussi de découvrir le monde professionnel et le monde du travail. Durant ces deux mois de stage effectués à INONI TECH, nous étions sous la coordination de deux encadrants, notamment un encadrant académique et un encadrant professionnel. Notre travail en tant que stagiaire était de concevoir un projet d'intelligence artificielle proposé par notre encadreur professionnel et validé par notre encadreur académique.

Ce projet est un mélange de deux domaines, la médecine et l'intelligence artificielle. Le thème est : **Détection du Cancer du sein : Conception d'une application d'intelligence artificielle basée sur le Deep Learning**

I. Contexte général de l'étude

Le cancer du sein est un cancer de la glande mammaire. Autrement dit, c'est un cancer qui naît dans les unités cellulaires dont la fonction est de sécréter le lait, les unités ducto-lobulaires du sein, essentiellement chez la femme. Depuis l'Antiquité, le cancer du sein a servi de modèle pour mieux connaître la maladie. C'est Hippocrate le premier qui l'a nommé crabe (karkinos en grec, cancer en latin).

Le cancer du sein est l'un des cancers les plus courants chez les femmes dans le monde entier, y compris en Afrique et en Europe. Selon l'OMS, l'Afrique détient le taux le plus élevé de mortalité par cancer du sein, avec 85800 femmes qui sont mortes en 2020, pour la majorité en Afrique de l'Ouest et en Afrique de l'Est. Les femmes qui meurent de cette maladie sont de plus en plus jeunes et se rendent à l'hôpital à un stade trop avancé. Celui-ci représente la première cause de décès par cancer chez les femmes africaines dans 34 pays du continent. Plusieurs facteurs expliquent ces mauvais chiffres en Afrique.

En Europe, on observe une augmentation régulière du nombre de nouveaux cas de cancer du sein depuis plusieurs décennies. En 2020, on estime que 355 457 nouveaux cas ont été diagnostiqués dans l'Union Européenne. Les tendances de l'incidence dans l'Union Européenne sont principalement en hausse. De multiples facteurs expliquent cette évolution, notamment les facteurs de risque, l'augmentation de la prévalence de l'obésité et de l'inactivité physique, ainsi que l'amélioration de la couverture de la population par le dépistage. Selon les données épidémiologiques les plus récentes, le taux de survie à cinq ans s'élève à plus de 90 % dans les pays à revenu élevé, mais n'atteint que 66 % en Inde et 40 % en Afrique du Sud. C'est en Afrique et en Polynésie que l'on observe le taux.

Ce projet constitue une aide dans la lutte contre le cancer du sein en faisant détecter les cas de cancer de sein grâce au Deep Learning

II. Problématique de l'étude

a. Présentation du problème

Le cancer du sein est un cancer qui se développe dans les cellules mammaires et progresse par étapes. Les cellules cancéreuses peuvent se propager du sein jusqu'à d'autres parties du corps, tels que l'aisselle, autour de la clavicule, du sternum. Il s'agit d'une maladie grave qui affecte un grand nombre de femmes à travers le monde. En 2021, il a été rapporté comme le cancer le plus courant chez les femmes, avec plus de 2,3 millions de nouveaux cas dans le monde. Malgré les progrès de dépistage et de traitement, le taux de mortalité reste élevé. La détection tardive, l'accès limité aux soins de santé, l'inégalité des soins et le manque de sensibilisation sont autant de facteurs contribuant à ce problème. Par conséquent, la recherche et l'innovation dans le domaine du dépistage précoce et du traitement du cancer du sein sont d'une importance cruciale pour sauver des vies et améliorer la qualité de vie des patientes.

b. Problème général

De ce qui précède, il convient de se poser la question suivante : Peut-on utiliser les techniques de l'intelligence artificielle pour permettre un dépistage rapide du cancer du sein ?

c. Problèmes spécifiques

Cette interrogation principale appelle aux deux préoccupations annexe suivantes :

- Peut-on améliorer le diagnostic du développement du cancer du sein au Cameroun à l'aide des outils de l'intelligence artificielle ? si oui quels en sont les résultats ?
- L'apprentissage en profondeur des données antérieures sur le cancer du sein peut-il nous permettre d'accélérer et de rendre un diagnostic précis ? si oui quels sont les résultats obtenus ?

III. Hypothèse de l'étude

a. Hypothèse générale

Dans le cadre de notre étude, nous pouvons poser l'hypothèse selon laquelle l'application de l'intelligence artificielle peut nous aider dans la lutte contre le cancer du sein par l'amélioration de la recherche, du dépistage, la détection précoce, le diagnostic.

b. Hypothèses spécifiques

Cette hypothèse générale peut nous conduire aux deux sous hypothèses suivantes :

- L'Intelligence Artificielle peut être utilisée pour analyser d'énormes volumes de données de recherche sur le cancer du sein, ce qui peut aider à découvrir de nouveaux traitements ou à améliorer les approches existantes.
- L'analyse des images médicales comme la mammographie de manière plus précise et rapide que les humains à l'aide de l'Intelligence Artificielle pourrait permettre la détection précoce du cancer du sein.
- En analysant de grand ensemble de données des patients atteints de la maladie, L'Intelligence Artificielle peut nous permettre de prédire l'évolution de la maladie chez tous les nouveaux cas.

IV. Objectif de l'étude

a. Objectif général l'étude

L'objectif de cette étude est de montrer l'importance de l'application de l'Intelligence Artificielle dans la lutte contre le cancer du sein

b. Objectif Spécifique de l'étude

L'objectif spécifique de cette recherche est de : Concevoir un modèle de classification d'image afin de détecter si une personne est atteinte ou non par le Cancer du sein. Evaluer et optimiser les résultats.

c. Justification de l'Etude

Prédire les nouveaux cas de cancer du sein à court, moyen et long terme permet aux autorités sanitaires d'instaurer des mesures préventives contre les impacts sur la santé publique et d'éviter des situations dramatiques similaires à celles rencontrées lors de pénuries médicales. Le diagnostic assisté par ordinateur est désormais un sujet de recherche essentiel dans l'analyse des images médicales et les diagnostics cliniques.

Les réseaux de neurones profonds ont connu des avancées notables dans l'identification d'anomalies, la segmentation sémantique et la classification d'images en matière de détection du cancer. Les tumeurs mammaires présentent des caractéristiques distinctes à l'imagerie, qui, bien que détectables par un radiologue, restent une tâche exigeante, pouvant conduire à des erreurs humaines. De plus, le temps nécessaire pour analyser le grand volume d'images peut être conséquent, d'où l'importance d'éviter toute confusion avec d'autres affections mammaires.

d. Délimitation de l'étude

Cette étude sera essentiellement limitée à la réalisation d'un algorithme pour la prédiction du cancer du sein, en utilisant les réseaux neuronaux convolutifs pour une classification des radiographies de mammographie.

e. Plan du mémoire

Outre l'introduction et la conclusion, cette étude se divise en quatre chapitres.

Le premier aborde le contexte conceptuel et théorique.

Le second décrit les méthodes employées pour le projet, incluant une définition des termes essentiels afin d'être dans un même esprit avec nos lecteurs et une présentation des données et outils.

Le troisième chapitre détaille l'entreprise où le stage a eu lieu et les informations recueillies.

Le dernier chapitre enfin expose les algorithmes développés et leurs retombées.



**CHAPITRE I : CADRE
THEORIQUE ET CONCEPTUEL**

Introduction

Le cadre théorique est un élément important de la recherche. Il permet non seulement de présenter les concepts clés, les théories, les idées préexistantes mais aussi de donner une justification scientifique à nos recherches en démontrant que celle-ci ne sortent pas de nulle part et qu'au contraire, elles possèdent de solides fondations scientifiques. L'Intelligence artificielle est aujourd'hui grandement utilisée dans le domaine de la médecine.

L'objectif de ce chapitre est d'une part de mettre en exergue les travaux des chercheurs qui ont essayé d'intégrer d'une manière ou d'une autre l'Intelligence Artificielle dans la lutte contre le cancer du sein et d'autre part de présenter l'historique de l'Intelligence Artificielle et son insertion dans le domaine de la médecine.

I. Cadre théorique et état d'art

L'Intelligence Artificielle (IA) est un domaine de recherche multidisciplinaire qui vise à créer des systèmes informatiques capables de simuler l'intelligence humaine. Elle a émergé en tant que discipline académique dans les années 1950, lorsque des chercheurs ont commencé à utiliser des machines pour simuler le raisonnement humain. L'IA a progressé lentement pendant plusieurs décennies, mais a commencé à prendre de l'ampleur à partir des années 1990 avec le développement d'algorithmes de Machine Learning plus sophistiqués et le progrès des technologies de l'information. En 2020, elle subit une constante évolution avec des avancées significative dans des domaines comme la génération de texte, les systèmes de recommandations personnalisés, la reconnaissance faciale, et plus encore.

L'IA repose sur plusieurs disciplines telles que l'informatique. Elle est constituée de plusieurs sous domaines notamment

- Le traitement du langage naturel
- La robotique
- La vision par ordinateur
- L'Intelligence Artificielle générale
- L'apprentissage automatique

L'Intelligence Artificielle a été introduite dans le domaine de la médecine dans les années 1970, initialement pour aider à l'interprétation des images médicales. Depuis lors, son application s'est étendue à de nombreux autres domaines, tels que la prédiction des maladies, l'aide au diagnostic, la personnalisation des traitements et l'amélioration de l'efficacité des soins de santé.

Dans le domaine de la sante, l'Intelligence Artificielle comprend des systèmes d'apprentissage profond pour l'interprétation des images médicales, des algorithmes de

prédiction pour anticiper l'évolution des maladies, et des systèmes d'aide à la décision pour aider les cliniciens à choisir les meilleurs traitements. Ces technologies ont été validées par des études cliniques et sont de plus en plus utilisées dans la pratique médicale. Même si L'IA n'a pas découvert le cancer du sein à proprement parler, car la maladie était déjà connue bien avant son arrivée, elle est maintenant utilisée de manière significative dans le domaine de la médecine pour aider à la détection et au diagnostic du cancer du sein.

Par exemple, Google Health en 2020, a développé un système d'intelligence artificielle qui a montré, dans des études, qu'il pourrait surpasser les radiologues humains dans la détection du cancer du sein à partir des mammographies, et améliorer les méthodes de détection et de diagnostic.

Constance Lehman, professeur de radiologie à la Harvard Medical School et directrice de la radiologie du sein au Massachusetts General Hospital, en collaboration avec Regina Barzilay, Professeur d'informatique au MIT, ont mené une étude publiée en 2020 qui utilisait l'IA pour prédire le risque de cancer du sein chez les femmes. Le modèle d'IA qu'elle et son équipe ont développé a été entraîné sur des données de plus de 100 000 patientes et s'est avéré plus précis que les méthodes traditionnelles de prédiction du risque.

Toujours dans l'objectif de lutter contre le cancer du sein, Dr. Humayun Irshad en 2016, alors qu'il travaillait à l'Institut Curie en France, lui et son équipe ont publié un article décrivant comment ils utilisaient l'IA pour aider à analyser les images histopathologiques du cancer du sein. Dr. Mazen Hassanain quant à lui en 2018, a dirigé une équipe à l'Université de Stanford qui a publié une recherche sur l'utilisation de l'IA pour améliorer la détection du cancer du sein dans les images mammographiques.

Parmi les revues littéraires sur la lutte contre le cancer du sein nous avons le livre "The Emperor of All Maladies : A Biography of Cancer" écrit par l'auteur et oncologue Siddhartha Mukherjee. Bien que ce livre ne se concentre pas uniquement sur le cancer du sein, il offre une perspective approfondie sur la maladie dans son ensemble. Par ailleurs, nous avons le Dr Susan Love reconnue pour son livre "Dr. Susan Love's Breast Book", qui est souvent cité comme la bible pour toute personne touchée par le cancer du sein car aborde tous les aspects de la maladie.

II. Historique

L'intelligence artificielle est une branche de l'informatique qui permet aux systèmes d'apprendre et d'exécuter des tâches normalement associées à l'intelligence humaine, telles que la reconnaissance vocale, la prise de décisions ou la perception visuelle. L'histoire de l'intelligence artificielle débute en 1943, avec la publication de l'article « A Logical Calculus of Ideas Immanent in Nervous Activity. » par Warren McCullough et Walter Pitts. Dans ce document, les scientifiques présentent le premier modèle mathématique pour la création d'un réseau de neurones. Le premier ordinateur à réseau de neurones, Snarc, sera créé en 1950 par deux étudiants de Harvard : Marvin Minsky et Dean Edmonds. La même année, Alan Turing publie le Turing Test qui sert encore aujourd'hui à évaluer les IA.

En 1952, Arthur Samuel crée un logiciel capable d'apprendre à jouer aux échecs de manière autonome. Le terme d'intelligence artificielle est prononcé pour la première fois durant la conférence « Dartmouth Summer Research Project on Artificial Intelligence. » de John McCarthy en 1956. Lors de cet événement, les chercheurs présentent les objectifs et la vision de l'IA. Beaucoup considèrent cette conférence comme la véritable naissance de l'intelligence artificielle telle qu'elle est connue aujourd'hui. En 1959, Arthur Samuel invente le terme de Machine Learning en travaillant chez IBM. De leur côté, John McCarthy et Marvin Minsky fondent le MIT Artificial Intelligence Project. En 1963, John McCarthy crée aussi le « AI Lab » de l'université de Stanford. Le Japon et les États-Unis investissent massivement dans la recherche en IA. Les entreprises dépensent plus d'un milliard de dollars par an dans les systèmes experts et l'industrie se développe. Malheureusement, le marché des machines « Lisp » s'effondre en 1987 face à l'apparition d'alternatives moins onéreuses. C'est le « deuxième hiver de l'IA ». Les entreprises perdent leur intérêt pour les systèmes experts. Les gouvernements américains et japonais abandonnent leurs projets de recherche, et des milliards de dollars ont été dépensés pour rien. Dix ans plus tard, les avancées technologiques permettent un renouveau de l'intelligence artificielle.

En 2008, Google réalise de formidables progrès dans le domaine de la reconnaissance de discours et lance cette fonctionnalité dans ses applications pour smartphones. En 2012, Andrew Ng nourrit un réseau de neurones à l'aide de 10 millions de vidéos YouTube en guise d'ensemble de données d'entraînement. Grâce au Deep Learning, ce réseau de neurones apprend à reconnaître un chat sans qu'on lui ait appris ce qu'est un chat. C'est le début d'une ère nouvelle pour le Deep Learning. Nouvelle victoire de l'IA sur l'Homme en 2016, avec la victoire du système AlphaGo de Google DeepMind sur Lee Sedol, le champion de jeu de Go. L'intelligence artificielle conquiert aussi le domaine du jeu vidéo, avec notamment DeepMind AlphaStar sur Star Craft ou OpenAI Five sur Dota 2. Le Deep Learning et le Machine Learning sont désormais utilisés par les entreprises de toutes les industries, pour une multitude d'applications. L'IA ne cesse de progresser et surprendre par ses performances. Le rêve d'une intelligence artificielle générale se rapproche de plus en plus de la réalité.

Les concepts, techniques et outils de l'intelligence artificielle sont utilisés dans les applications médicales depuis plus de quatre décennies. L'utilisation de l'IA a été développée dans la perspective d'améliorer les soins en aidant les professionnels de la santé à améliorer leur efficacité, leur productivité et leur constance dans la qualité des soins apportés aux patients. L'amélioration de la précision et de l'efficacité des techniques d'IA n'a cessé de s'accroître et de s'intégrer dans les outils destinés à aider les soignants confrontés à des problèmes de plus en plus complexes. Les progrès technologiques ont contribué à rendre disponibles des algorithmes et ont permis l'adoption de l'IA pour de nombreuses applications médicales. Ces algorithmes d'IA ont été appliqués avec succès et couvrent un large éventail de domaines de la médecine. Ils concernent des applications médicales traditionnelles comme l'aide au diagnostic, à la décision thérapeutique et à l'imagerie médicale. Historiquement, l'idée que les ordinateurs pourraient aider à évaluer les probabilités diagnostiques alternatives n'est pas neuve et remonte aux années 1960. Les premières applications ont vu le jour concrètement au début des années 1970 en réponse à l'augmentation de la demande de services médicaux de qualité et à l'accroissement des connaissances médicales. Pour exemple, les premiers travaux sur les applications médicales de l'IA, portaient sur le diagnostic automatisé du dysfonctionnement thyroïdien

Ainsi, tous les secteurs d'activité ont été pratiquement envahis par l'IA, à savoir le transport, la santé, l'industrie manufacturière, la finance, les assurances, l'agriculture, les énergies classiques et renouvelables, la distribution, les médias, la justice, le tourisme, l'éducation, les services publics, la défense et les renseignements.

III. Cadre Règlementaire

L'Intelligence Artificielle est un domaine complexe et en évolution constante, qui varie grandement d'un pays à l'autre. Il est également en cours d'élaboration, car les gouvernements du monde entier cherchent à trouver un équilibre entre la promotion de l'innovation technologique et la protection des droits individuels et de la sécurité publique.

Au Cameroun, il n'existe pour l'heure aucune législation conçue pour réglementer l'utilisation de l'IA. Au contraire, les systèmes d'IA sont régulés par d'autres réglementations existantes mais non spécifiques. Cela inclut les lois sur la protection des données. Par exemple, les dispositions légales sur la protection des données se trouvent dans plusieurs lois. Comme les lois spécifiques sur la protection des données n'ont pas encore été adoptées, il est difficile pour les utilisateurs de contrôler l'utilisation de leurs données. La législation applicable couvre principalement les données de communication électronique, tandis que d'autres industries traitent quotidiennement des données personnelles.

L'Article 61 de la loi sur la cybersécurité prévoit un certain nombre de sanctions en cas de violation des données. En outre, l'article 74 de la loi sur la cybersécurité prévoit une peine d'emprisonnement d'un à deux ans et une amende de 1000000 XAF (environ 1500 €) à 5000000 (environ 7600 €) à quiconque porte atteinte à la vie privée d'autrui en fixant, enregistrant ou transmettant, sans le consentement de l'auteur, des données électroniques à caractère privé ou confidentiel. Il n'existe aucune disposition spécifique imposant des limites aux transferts de données. Lorsqu'un règlement à l'amiable ne peut être trouvé entre les parties concernées, les parties doivent se référer à l'ANTIC. Si la décision rendue par l'ANTIC n'est pas satisfaisante, les parties peuvent alors demander réparation devant les tribunaux, article rédigé par Danielle Moukouri, Managing Partner du cabinet camerounais membre Loi Moukouri

Il est important de noter que la réglementation de l'IA est un domaine en évolution rapide et que de nombreux pays travaillent actuellement à l'élaboration de leurs propres cadres réglementaires pour l'IA. En outre, il existe des discussions internationales sur la création de normes mondiales pour la réglementation de l'IA.

Conclusion

Dans ce premier chapitre, nous avons approfondi le sujet de l'intelligence artificielle et son application dans le secteur médical. Nous avons également pris en compte les résultats des recherches précédentes pour éclairer notre propre sujet. Pour ce faire, nous nous sommes appuyés sur divers articles et travaux académiques axés sur la lutte contre le cancer du sein. Dans le chapitre suivant, nous décrirons en détail les méthodologies employées dans notre recherche, après avoir décrit le type de recherche effectuée.



**CHAPITRE II : Méthodologie de
recherche**

Introduction

Tout travail scientifique demande l'emploi de méthodologies et techniques afin d'obtenir un résultat précis et efficace, en adéquation avec le sujet de l'étude. On peut définir la méthodologie comme l'ensemble de procédures, règles et méthodes qui facilitent le choix d'outils statistiques appropriés pour l'analyse de données.

Elle donne la possibilité à l'analyste de vérifier la qualité de ses travaux et d'atteindre ses buts. Dans ce contexte, nous avons utilisé diverses techniques et méthodes pour parvenir au résultat souhaité. Ce chapitre débute par la présentation de l'essence de notre recherche et des variables liées à notre questionnement. Par la suite, nous introduirons certains concepts liés à notre travail, tels que la modélisation, l'apprentissage profond (Deep Learning), et les méthodes d'intelligence artificielle. Finalement, nous mentionnerons non seulement les méthodes et techniques employées pour effectuer cette recherche, mais également les outils utilisés.

I. Nature de recherche

La nature de l'étude abordée dans ce document est perceptuelle. Le but est d'explorer et de décrire les données recueillies, d'élucider les concepts pour une meilleure compréhension du thème de recherche et d'anticiper les résultats.

II. Variable de recherche

Les variables utilisées pour la réalisation pour la réalisation de cette recherche sont les images de mammographies.

a. Définition conceptuelle

Un modèle informatique est une représentation mathématique ou algorithmique qui imite une certaine tâche ou un certain comportement. Il est construit à partir d'un ensemble de règles, d'équations, ou en utilisant des données d'apprentissage dans le cas de l'apprentissage automatique.

Machine Learning (apprentissage automatique) : est un sous-domaine de l'intelligence artificielle. Il se focalise sur la conception de systèmes qui peuvent apprendre de l'expérience, c'est-à-dire des données.

Le but de l'apprentissage automatique est de permettre aux machines d'apprendre automatiquement sans intervention humaine ou programmation explicite. Les algorithmes d'apprentissage automatique s'améliorent avec l'expérience, c'est-à-dire avec l'exposition à davantage de données au fil du temps.

Les algorithmes d'apprentissage peuvent se catégoriser selon le mode d'apprentissage qu'ils emploient. La figure suivante (figure 1) montre la classification des différents algorithmes du Machine Learning.

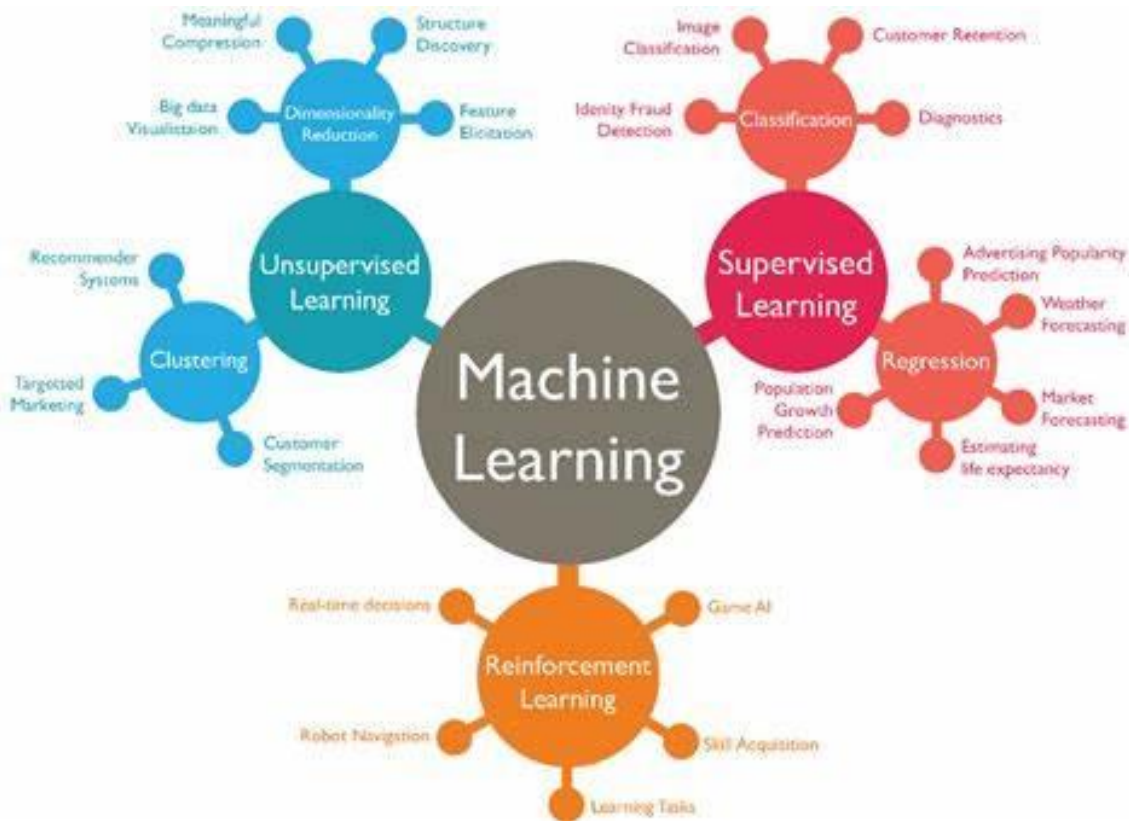


FIGURE 1 : ALGORITHME DE MACHINE LEARNING

On peut réorganiser les concepts en trois catégories d'algorithmes pour l'apprentissage automatique :

- **L'apprentissage supervisé** est un procédé où un modèle est formé en utilisant des données pré-annotées pour prédire des résultats. On divise cela en deux parties : la régression pour prédire des valeurs quantitatives, et la classification pour prédire des valeurs qualitatives.
- **L'apprentissage non supervisé** se base sur des données non labellisées, son objectif est de révéler les structures inconnues dans ces données. Malgré son imprévisibilité et l'impossibilité de mesurer son succès de manière certaine, il offre la capacité de résoudre des problèmes plus complexes comparé à l'apprentissage supervisé.
- **L'apprentissage par renforcement** est une approche où l'algorithme apprend de ses actions en fonction de leur impact sur l'environnement. Le Q-Learning est un exemple typique de cet apprentissage.

Enfin, il y a le **Deep Learning** consiste en des techniques d'apprentissage automatique qui visent à représenter des données à un niveau d'abstraction élevé à travers des séquences de transformations non linéaires.

Ces méthodes ont conduit à des avancées notables, notamment dans l'analyse sonore et visuelle, la reconnaissance faciale, la reconnaissance vocale, la vision par ordinateur et le traitement automatique du langage.

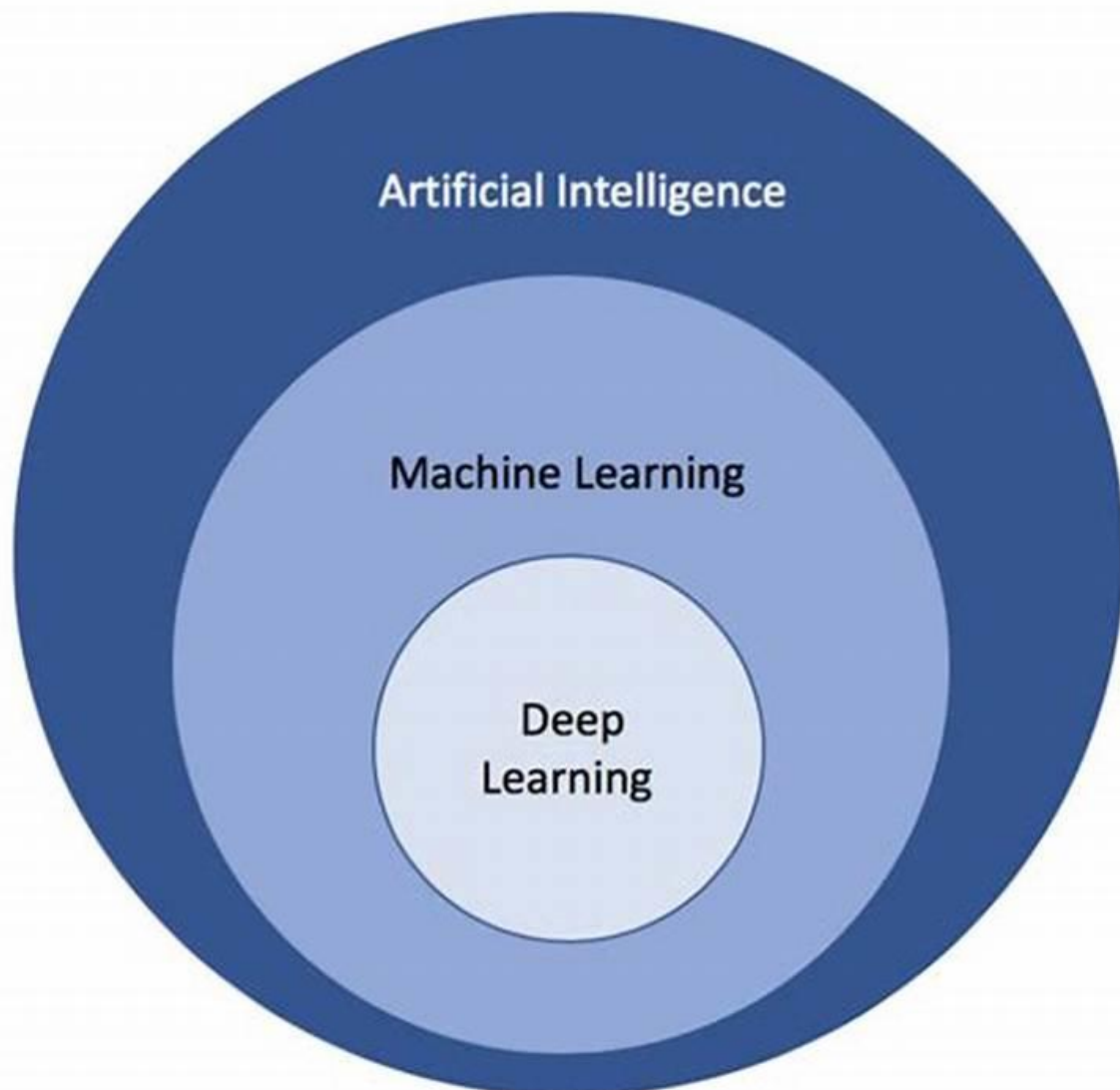
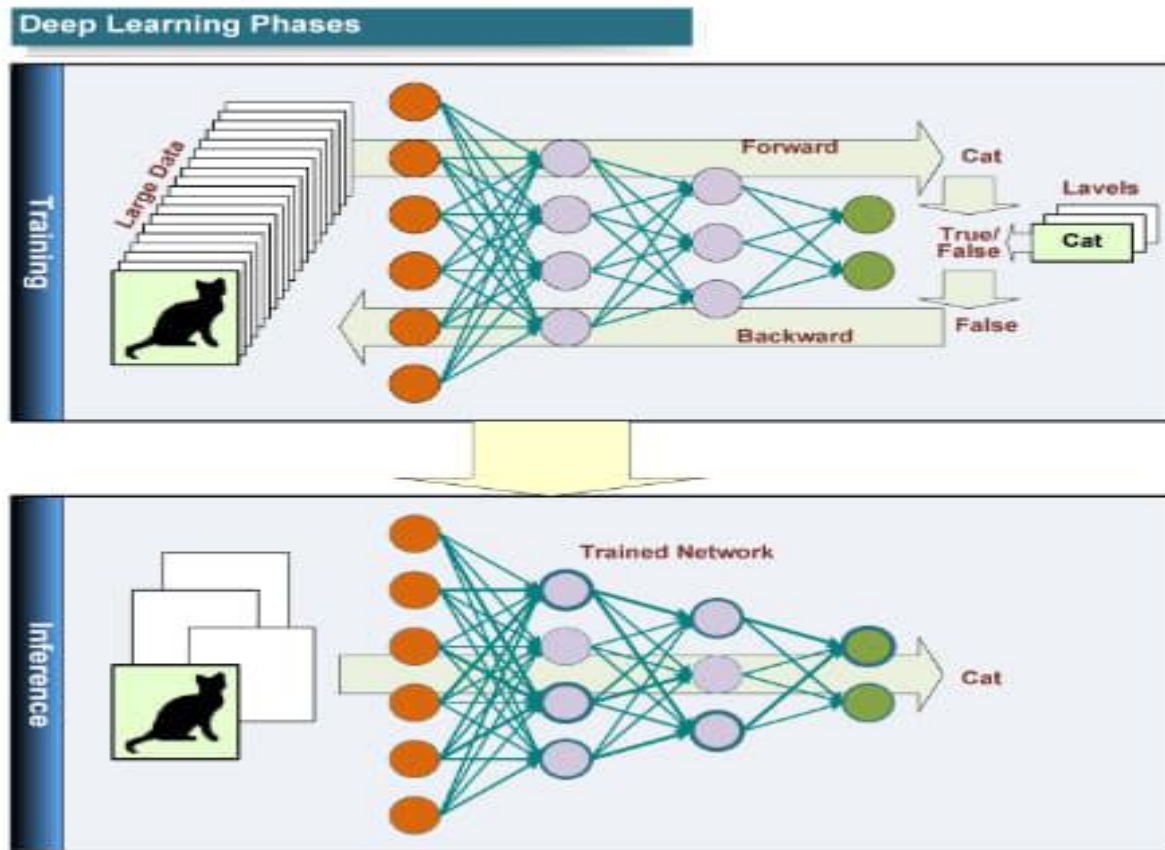


Figure 2 : Intelligence Artificielle et ses Composants

Les réseaux neuronaux convolutifs (CNN) sont une spécialité des réseaux de neurones destinés à traiter des images. Leur structure comprend quatre étapes principales :

- **La couche de convolution** : elle scrute les images entrantes pour identifier certaines caractéristiques ou "features".
- **La couche de pooling** : il compacte l'image en conservant ses informations cruciales.
- **La couche d'activation ReLU (Rectified Linear Units)** : elle élimine les valeurs négatives, introduisant ainsi de la non-linéarité, rendant le modèle plus élaboré.
- **Couche Fully Connected (FC)** : elle est liée à tous les neurones de sortie et a pour mission de classer l'image entrante en fonction des probabilités associées à chaque catégorie.



Copyright (c) 2017 Hiroshige Goto All rights reserved.

FIGURE 3 : ARCHITECTURE D'UN CNN

Instruments de recherche

Microsoft power BI  :

Est une solution d'analyse de données de Microsoft. Il permet de créer des visualisations de données personnalisées et interactives avec une interface suffisamment simple pour que les utilisateurs finaux créent leurs propres rapports et tableaux de bord.

Python  :

Est un langage de programmation interprété, interactif et orienté objet, qui possède une sémantique dynamique. Grâce à une multitude de bibliothèques spécialisées, il peut être utilisé dans divers contextes et s'adapter à toutes sortes d'utilisations. Il a été le langage de choix pour la réalisation de ce projet.



Anaconda :

Est une distribution gratuite et open source des langages de programmation Python et R qui est utilisée pour développer des applications dédiées à la science des données et à l'apprentissage automatique. Elle simplifie la gestion des paquets et le déploiement.



Jupyter Notebook

Est un environnement de calcul interactif en ligne permettant de créer des documents de bloc-notes. Construit avec plusieurs bibliothèques open source, il gère les versions de package dans Anaconda grâce au système de gestion de package Conda.



La bibliothèque Numpy

Est un outil pour Python, spécialement conçu pour manipuler des matrices ou des tableaux multidimensionnels et les fonctions mathématiques qui les opèrent.

Pandas

Est une autre bibliothèque Python qui facilite la manipulation et l'analyse des données.



Matplotlib

Est une bibliothèque Python qui permet de tracer et visualiser des données sous forme de graphiques.



Scikit-Learn

Est une bibliothèque Python dédiée à l'apprentissage automatique, incluant des fonctions pour l'estimation de régressions logistiques, d'algorithmes de classification, et de machines à vecteurs de support.



Tensorflow

Est une bibliothèque logicielle open source destinée au calcul numérique de haute performance. Créée par Google, elle vise à faciliter la création de modules d'apprentissage automatique et est particulièrement adaptée pour les réseaux de neurones, même si elle est également utilisable pour d'autres calculs modélisables sous forme de graphiques de flux de données.



Keras

Est une bibliothèque qui facilite l'interaction avec les algorithmes de réseaux de neurones profonds et d'apprentissage automatique, comme Tensorflow. Elle est conçue pour permettre une expérimentation rapide avec les réseaux de neurones profonds.

Conclusion

Ce chapitre a présenté la nature de notre étude, et a brièvement introduit les variables de recherche liées à notre travail en définissant leurs concepts. Nous avons également présenté les différents outils utilisés pour la réalisation de notre projet.



**CHAPITRE III : PRESENTATION
D'INONI TECH ET DES DONNEES
COLLECTEES**

Introduction

Dans ce chapitre nous présenterons la structure dans laquelle nous avons effectué ce stage académique tant dans le cadre administratif que technique. Dans le cadre administratif, nous présenterons la hiérarchie l'entreprise INONI TECH son activité et sa localisation, dans le cadre technique, nous présenterons les tâches que nous avons effectués durant cette période de stage.

I. Présentation de l'entreprise

Inoni Tech est le premier pôle technologique de la région dédié au Web 3 et à la technologie blockchain. Il s'engage à construire un écosystème dynamique et collaboratif dans lequel les entrepreneurs sociaux, les startups tech et les investisseurs d'impact peuvent travailler ensemble pour libérer le potentiel inexploité et créer des opportunités pour l'avenir.

a. Produits et services de l'entreprise

Pour atteindre leurs objectifs fixés à court, moyen et long terme, ils offrent une panoplie de produits/services à la communauté :

- Des formations pratiques sur l'écosystème de la blockchain, web 3.0, les cryptomonnaies : Comprendre comment fonctionne cette technologie est fondamental pour s'y employer définitivement et contribuer significativement au développement du tissu économique national et impacter une plus grande communauté.
- Des salles de réunions
- Des bureaux collaboratifs (coworking space)
- Des salles de formations
- Un vaste espace outdoor pour des rencontres professionnelles, des événements d'entreprise, des espaces de travail, un espace calme. Nos bureaux vous appartiennent.
- Des formations en marketing digital et Community management
- Des formations en infographie et montage audiovisuel.

b. Localisation de l'entreprise

IMOMI TECH est situé à bastos en face de l'ancien ambassade d'Italie. Les plans de localisation ci-dessous sont tirés de Google Maps.

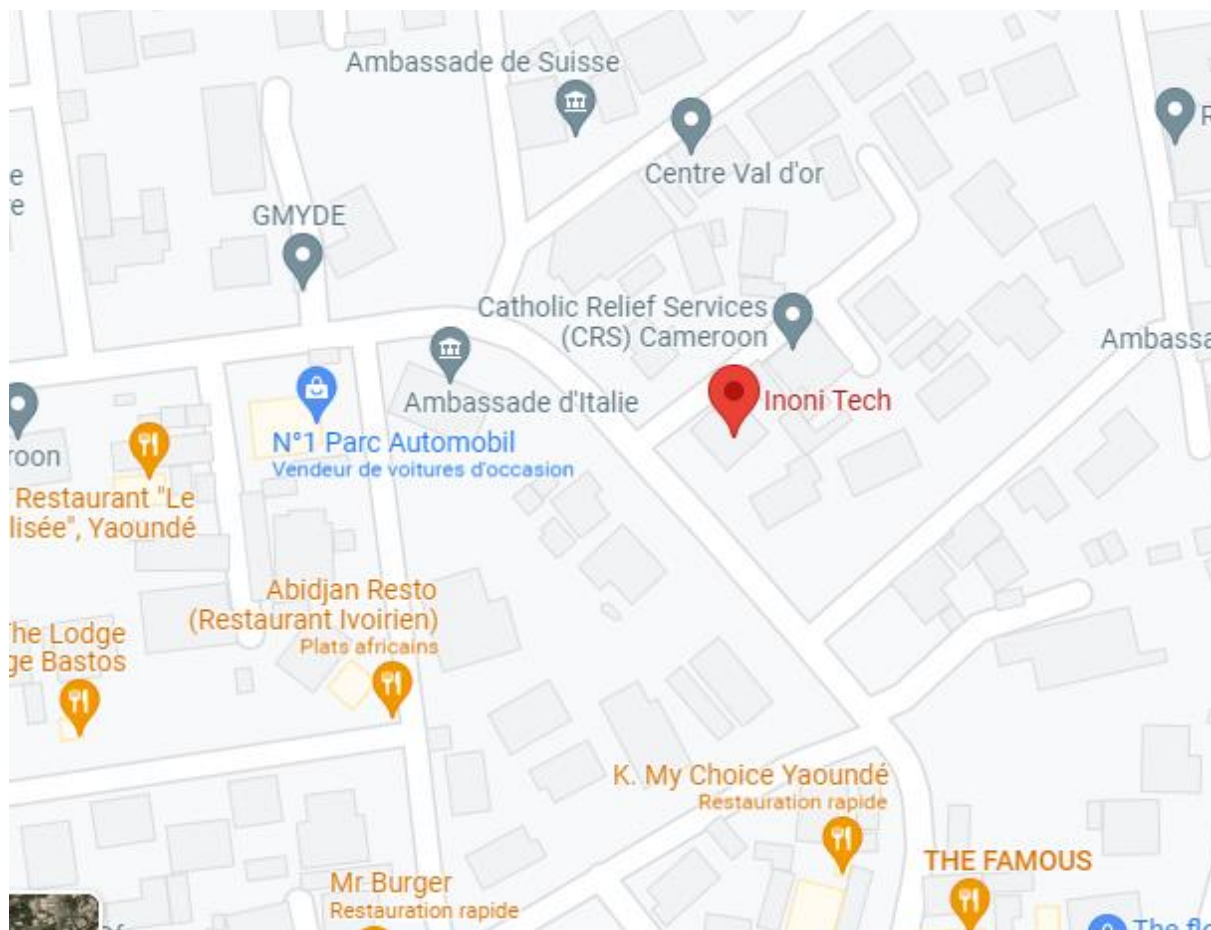


FIGURE 4 : LOCALISATION DE INONI TECH

c. Mission

Notre mission est d'habiliter les innovateurs sociaux, les entreprises tech et les investisseurs d'impact à cocréer des solutions qui changeront le continent africain pour le mieux.

C'est le premier centre éducatif en Afrique francophone orienté web 3.0, technologie blockchain ; cryptoactifs. En partenariat avec la plus grande plateforme d'échange des cryptomonnaies au monde, Binance qui nous appuie en offrant des formations gratuites aux apprenants et aux amateurs de technologie blockchain. Nous mettons à la disposition du grand public et des entrepreneurs sociaux et tech nos locaux constitués de plusieurs salles de réunions, un espace de formation, un espace de travail collaboratif pour favoriser le développement et l'implémentation d'idées innovantes. Nous sommes un incubateur. Par ailleurs, nous offrons des formations aux métiers actuels et d'avenir aux jeunes pour leur offrir de devenir des professionnels dans les domaines choisis et aux entrepreneurs pour les aider à améliorer leurs savoirs pratiques et leurs prodiguons des conseils stratégiques, en plus du réseautage, gage de pérennisation de leurs affaires.

d. Organigramme de l'entreprise

Inoni Tech est organisé de la façon suivante :

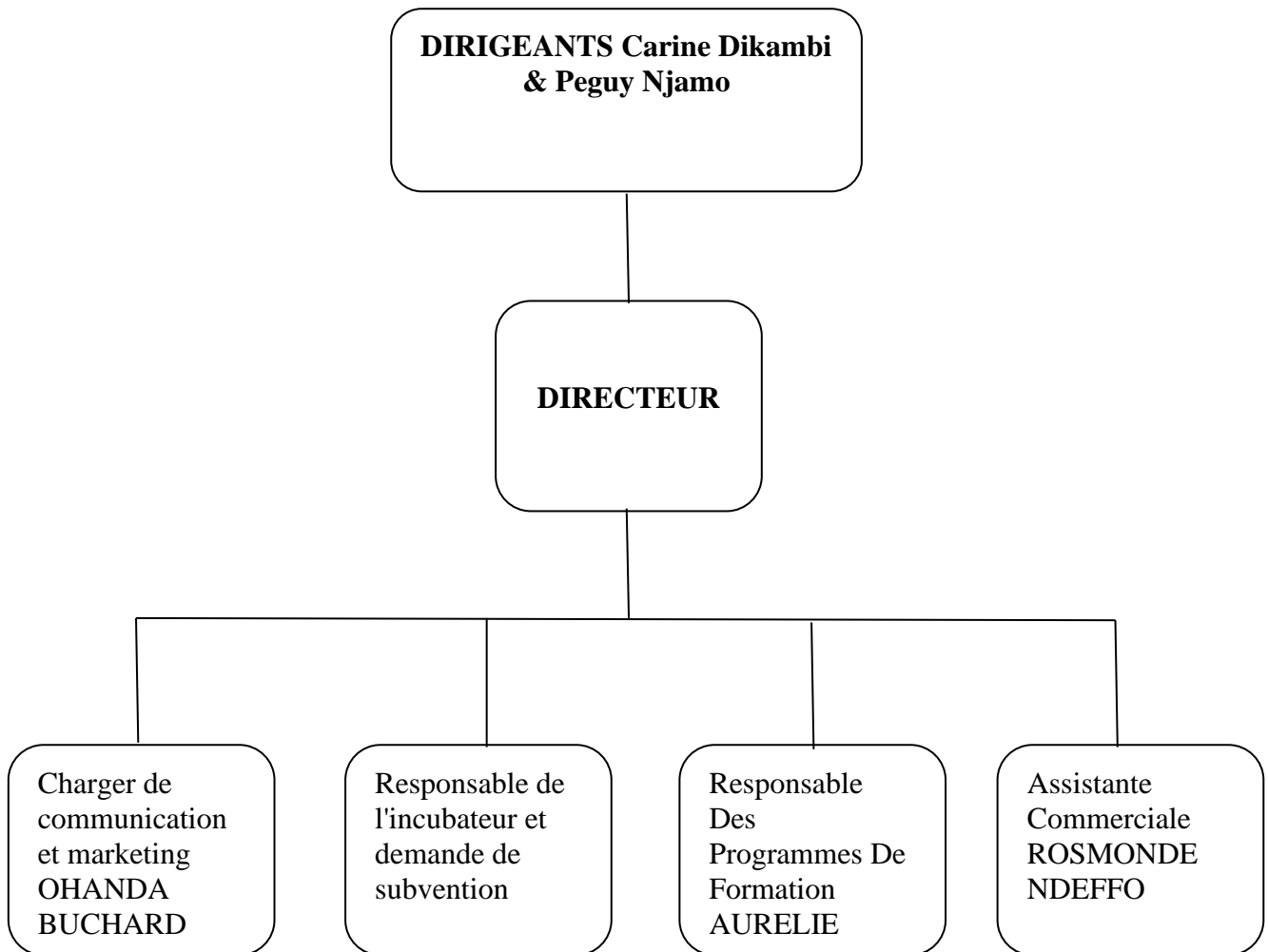


FIGURE 5 : ORGANIGRAMME DE L'ENTREPRISE

II. Données collectées

Lors de notre stage chez INONI TECH, nous avons travaillé sur diverses tâches pour approfondir nos connaissances en Big Data. Parmi elles, nous avons étudié et produit des Dashboards sur Tableau, étudié les étapes de réalisation d'un chatbot. Suite à cela, nous avons démarré notre projet principal. Pour ce projet, nous avons utilisé un ensemble de données qui contient des informations sur la progression du cancer du sein telles que les cas confirmés, les personnes atteintes de relations bilatérales mastectomie.

Pour notre algorithme de Deep Learning, nous avons utilisé des images de mammographie provenant de différentes sources, y compris des hôpitaux. Ces images, disponibles publiquement sur Kaggle, montrent des patients atteints du Cancer du sein, des patients guéris

et ceux normaux. L'ensemble comprend 1400 images, réparties en deux catégories, dont 980 pour l'entraînement et 420 pour les tests.

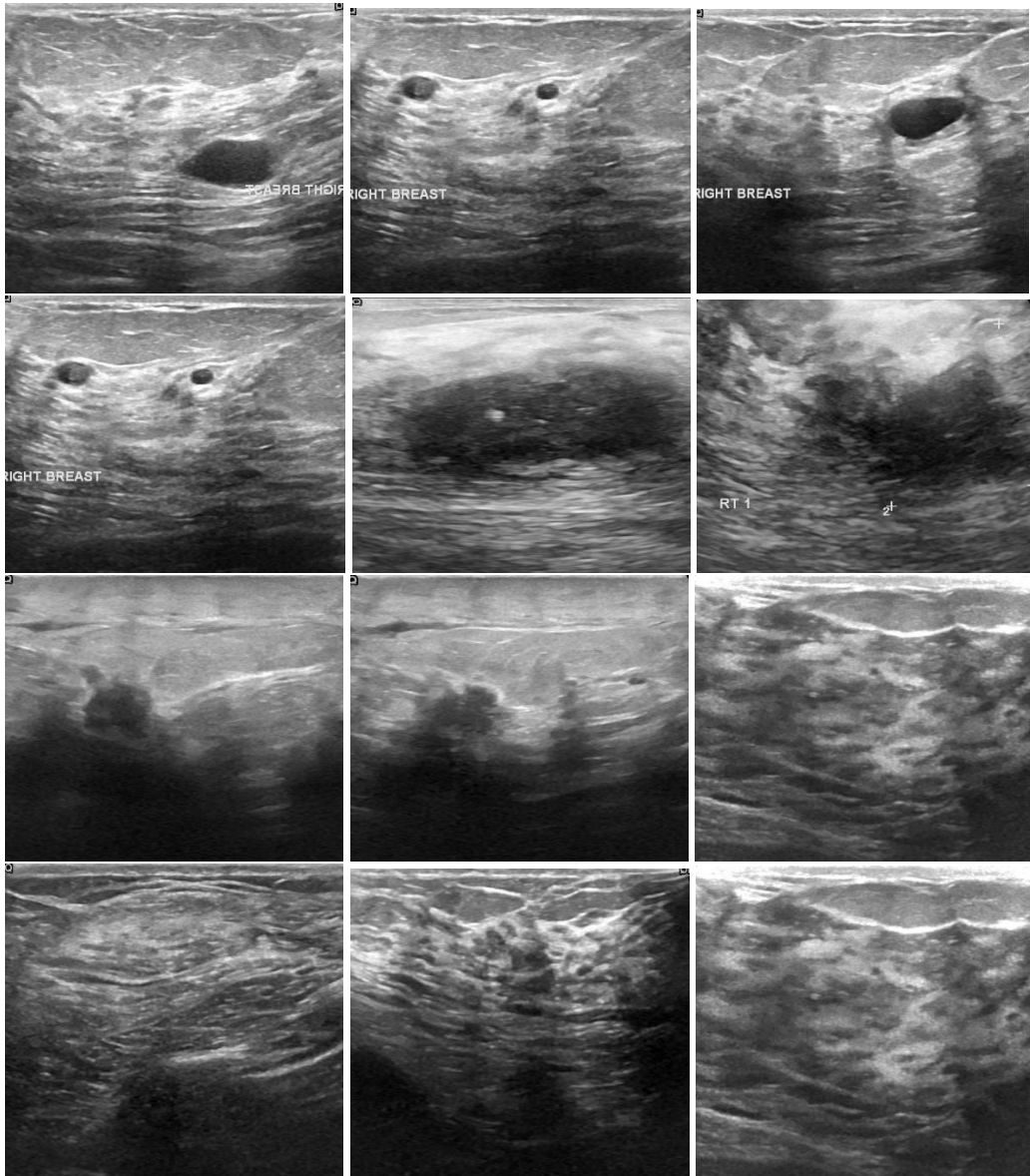


FIGURE 6 : MAMMOGRAPHIE

Conclusion

Ce chapitre décrit INONI TECH, l'entreprise où nous avons réalisé notre stage. Nous y détaillons son organisation, ses services et ses objectifs principaux. Ensuite, nous évoquons les données que nous avons rassemblées pour notre mémoire.



**CHAPITRE IV : ANALYSE
DIAGNOSTIQUE DE LA
SITUATION ET PROPOSITION
D'INTERVENTION**

Introduction

Dans ce chapitre, nous introduirons les tableaux de bord créés, suivis d'une analyse courte de ceux-ci. Nous explorerons également des exemples de code pour le Deep Learning. Les démarches pour développer nos modèles d'expérimentation en DL, seront détaillées, tout comme les tests réalisés. En conclusion, nous comparerons et discuterons des performances du modèle.

I. Présentation et analyse de la situation

Le cancer du sein a touché des millions de personnes dans le monde au fil des années. Sa nature agressive et ses diverses formes le rendent particulièrement difficile à traiter. L'Organisation mondiale de la santé (OMS) a reconnu le cancer du sein comme l'un des cancers les plus diagnostiqués chez les femmes. Il est difficile de détecter précocement la maladie, car elle peut se développer sans symptômes apparents pendant une longue période. Depuis les premiers recensements, le nombre de cas a augmenté dans presque tous les pays, entraînant le décès de millions de personnes. La figure ci-dessous illustre l'ampleur du cancer du sein dans le monde à ce jour.

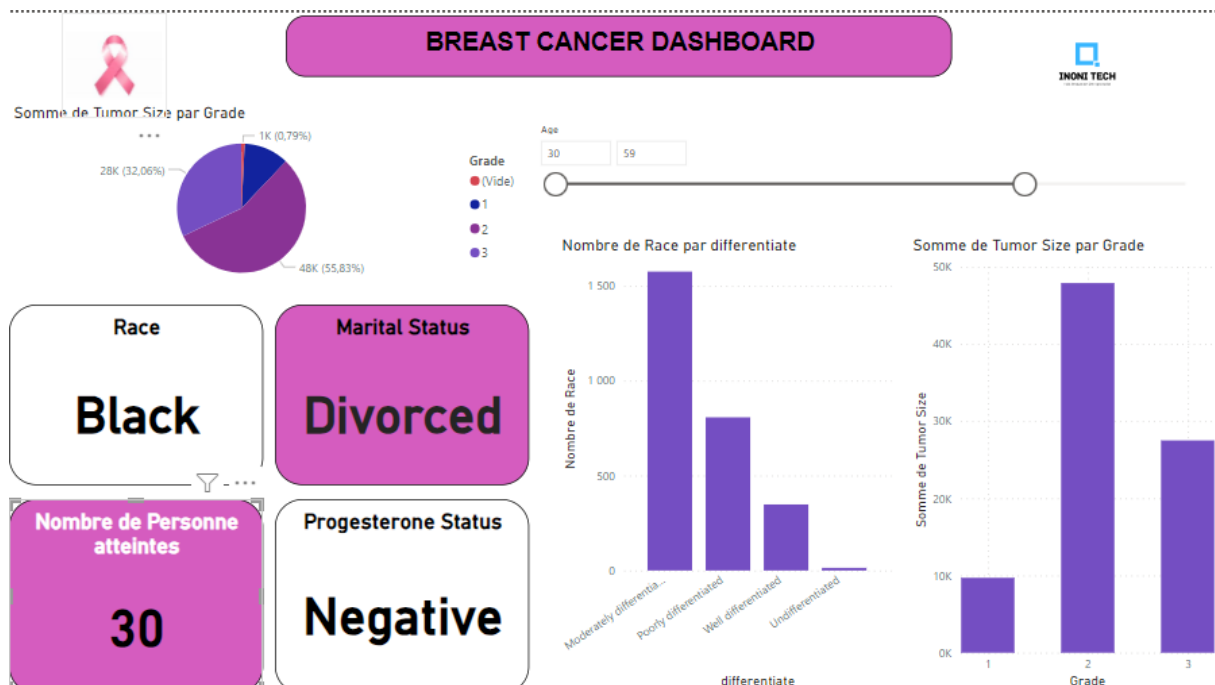


FIGURE 7 : PRESENTATION DE L'ÉVOLUTION DU CANCER DU SEIN

Analyse :

Le tableau de bord ci-dessus vise à fournir une vue d'ensemble des caractéristiques et des facteurs de risque associés au cancer du sein, ce qui peut être essentiel pour la recherche, le diagnostic et le traitement. Nous avons

- ✓ L'Age qui est un facteur clé, car le risque de développer un cancer du sein augmente avec l'âge comme nous pouvons le constater grâce au Dashboard
- ✓ Statut marital : Peut-être important pour comprendre les facteurs sociaux ou de soutien qui pourraient influencer les résultats.
- ✓ Taille de la tumeur : Un facteur clé dans le diagnostic et le pronostic du cancer, qui est illustré par un histogramme qui montre la distribution des tailles de tumeurs.
- ✓ Grade de la tumeur : Il indique la rapidité de croissance et la probabilité de propagation de la tumeur.
- ✓ Statut d'estrogènes et de progestérone : Ces récepteurs hormonaux jouent un rôle crucial dans le traitement du cancer du sein.

La première étape du traitement du cancer du sein consiste à dépister les patientes dans les centres de santé primaires ou les hôpitaux. Bien que le diagnostic final repose encore principalement sur des biopsies et des analyses histologiques, l'imagerie médicale joue un rôle crucial pour évaluer l'extension de la maladie. Les patientes suspectées d'avoir un cancer du sein subissent généralement une mammographie, et si plus de détails sont nécessaires, elles peuvent avoir recours à une échographie ou une IRM mammaire.

Grâce à ces techniques, la mammographie et l'échographie sont largement utilisées en clinique comme outils de diagnostic primaires pour détecter le cancer du sein. Pour établir un diagnostic, les médecins examinent les images à la recherche de masses ou de calcifications suspectes. La prévalence élevée du cancer du sein conduit à un grand nombre de patientes nécessitant un diagnostic, ce qui peut entraîner une surcharge pour les radiologues et une éventuelle pénurie de spécialistes en oncologie.

II. Intervention proposée et justification

L'analyse prédictive du cancer du sein est devenue un domaine de recherche essentiel pour soutenir les cliniques oncologiques et les systèmes de santé, afin de planifier et d'optimiser les traitements. La modélisation et la prévision du taux d'incidence du cancer du sein peuvent aider les systèmes de santé à être prêts à fournir les traitements nécessaires en temps voulu. Une prévision précise de l'évolution de cette maladie est cruciale car elle peut influencer les décisions politiques, les protocoles médicaux, les systèmes de santé et la qualité de vie des patients. Dans notre étude, nous nous concentrerons sur un modèle de Deep Learning basé sur TensorFlow pour détecter les cas de cancer du sein à un stade avancé.

III. Objectifs de l'intervention - projet envisagé

a. Objectif général

L'objectif de cette étude est de montrer l'importance d'Appliquer l'Intelligence Artificielle dans la lutte contre le cancer du sein

b. Objectifs spécifiques

En 2021, plus de 17 millions de nouveaux cas de cancer du sein ont été diagnostiqués, avec 671 500 décès associés. Malgré les avancées médicales, il n'existe pas encore de solution définitive pour contrer cette maladie dévastatrice. Notre objectif est d'établir un modèle de classification en Deep Learning capable de déterminer la probabilité qu'une personne soit atteinte ou non du cancer du sein.

IV. Composante de l'intervention envisagée

Cette section est principalement une analyse de l'état actuel, qui vise à comprendre comment intégrer des améliorations durables dans le domaine médical grâce à l'intelligence artificielle. L'étude se focalise sur la classification des personnes atteintes de cancer du sein, et celles en bonne santé.

a. Stratégies d'action et contenu : Conception d'un modèle Deep Learning pour la détection du Cancer du sein

Il s'agit d'un code classique de construction, entraînement et évaluation d'un modèle de classification binaire utilisant un réseau de neurones convolutifs (CNN) avec TensorFlow et Keras. C'est un modèle de Deep Learning pour classer des images comme "malignes" ou "normales".

Les images en entrée ont une taille de 32x32, chaque couche de convolution est composée de filtres. Dans notre travail nous avons utilisé 32 filtres pour les deux premières couches. Après l'étape de convolution, 32 plans (ou cartes) de caractéristiques (features maps) de taille 64x64 seront créés

Au début, on importe diverses bibliothèques nécessaires pour le traitement des données, la visualisation et le Deep Learning, puis on initialise ensuite un modèle séquentiel appelé "classifier", ce qui signifie que les couches sont ajoutées les unes après les autres. On ajoute deux couches convolutionnelles pour extraire des caractéristiques des images. Ces couches sont suivies de couches de pooling pour réduire la taille des représentations spatiales et ainsi améliorer l'efficacité.

```

classifler = Sequential()

# Step 1 - Convolution
classifler.add(Conv2D(32, (3, 3), input_shape = (64, 64, 3), activation = 'relu', padding= 'same'))

# Step 2 - Pooling
classifler.add(MaxPooling2D(pool_size = (2, 2)))

# Adding a second convolutional layer
classifler.add(Conv2D(32,3,3, activation = 'relu', padding= 'same'))
classifler.add(MaxPooling2D(pool_size = (2, 2)))

# Step 3 - Flattening
classifler.add(Flatten())

# Step 4 - Full connection
classifler.add(Dense(units = 128, activation = 'relu'))
classifler.add(Dense(units = 1, activation = 'sigmoid'))

# Full connection
#Early Stopping
callback = tf.keras.callbacks.EarlyStopping(monitor='val_loss', mode='min', verbose=1,patience=2)

```

FIGURE 8 : CREATION ET OPTIMISATION DU MODELE

La couche "Flatten" est utilisée pour transformer la matrice 2D en un vecteur 1D. Deux couches denses (ou fully connected) sont ajoutées. La dernière couche utilise une activation "sigmoid", ce qui signifie que le modèle prédira la probabilité qu'une image soit "maligne" ou "normale".

Ensuite, trois callbacks sont définis : **EarlyStopping** est une forme de régularisation utilisée pour éviter le surapprentissage lors de la formation d'un apprenant avec une méthode itérative, telle que la descente de gradient. Elle permet d'arrêter l'entraînement lorsqu'une métrique surveillée a cessé de s'améliorer. Dropout : Le problème de surapprentissage est généralement engendré lorsque certains neurones sont trop liés entre eux et trop dépendants et ce, au fur et à mesure de l'apprentissage. **ModelCheckpoint** pour sauvegarder le meilleur modèle basé sur la précision de validation. **ReduceLRonPlateau** pour réduire le taux d'apprentissage si la performance ne s'améliore pas.

Le code définit ensuite des générateurs d'images pour augmenter et normaliser les images d'entraînement et de validation. Le modèle est formé sur l'ensemble d'entraînement pendant 15 époques. La structure du modèle est affichée grâce à **classifler.summary()**.

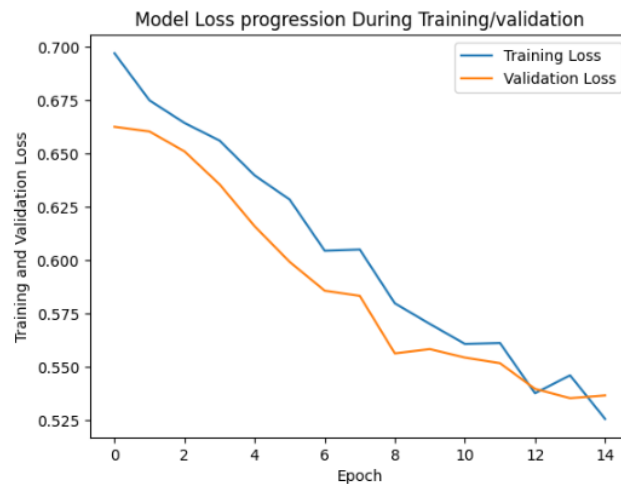
```
Entrée [17]: classfier.summary()
Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 64, 64, 32)         896
max_pooling2d (MaxPooling2D) (None, 32, 32, 32)         0
conv2d_1 (Conv2D)            (None, 11, 11, 32)         9248
max_pooling2d_1 (MaxPooling2D) (None, 5, 5, 32)         0
flatten (Flatten)            (None, 800)                 0
dense (Dense)                 (None, 128)                102528
dense_1 (Dense)               (None, 1)                  129
-----
Total params: 112801 (440.63 KB)
Trainable params: 112801 (440.63 KB)
Non-trainable params: 0 (0.00 Byte)
-----
```

FIGURE 9 : STRUCTURE DU MODELE

Enfin, une image test est chargée et prédite par le modèle. Si la probabilité prédite est 0, l'image est classée comme "maligne", sinon elle est classée comme "normale".

b. Résultats obtenus et discussion

Les résultats d'entraînement et du test obtenus à partir de notre modèle de classification DL pour une période de 15 époques (une époque, décrit le nombre de fois que l'algorithme passe sur l'ensemble de données) sont présentés sur la Figure ci-dessous).



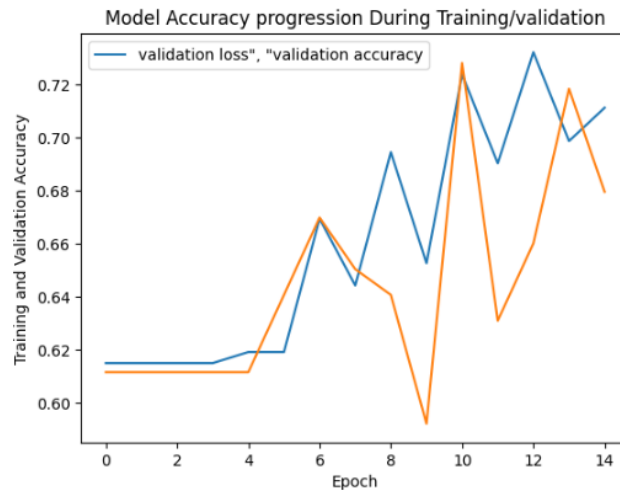


FIGURE 10 : RESULTAT DU MODELE

Ces courbes de perte et de précision pour l'entraînement et la validation sont tracées pour aider à visualiser la performance du modèle au fil des époques.

Les résultats d'entraînement de notre modèle de classification Deep-Learning pour une durée de 15 époques montrent que les courbes de l'entraînement et de validation diminuent toutes les deux pour l'erreur (Loss) et elles augmentent ensemble pour la précision (Accuracy). Elles évoluent ensemble vers le bon sens et tendent vers de meilleurs résultats malgré l'existence de certaines oscillations dans les courbes de validation mais ceci ne dégrade pas la performance du modèle.

```

- - - - -
Epoch 8/15
8/8 [=====] - 6s 752ms/step - loss: 0.6051 - accuracy: 0.6444 - val_loss: 0.5834 - val_accuracy: 0.650
5 - lr: 0.0010
Epoch 9/15
8/8 [=====] - 6s 811ms/step - loss: 0.5799 - accuracy: 0.6946 - val_loss: 0.5564 - val_accuracy: 0.640
8 - lr: 0.0010
Epoch 10/15
8/8 [=====] - 6s 769ms/step - loss: 0.5702 - accuracy: 0.6527 - val_loss: 0.5585 - val_accuracy: 0.592
2 - lr: 0.0010
Epoch 11/15
8/8 [=====] - 6s 822ms/step - loss: 0.5609 - accuracy: 0.7238 - val_loss: 0.5545 - val_accuracy: 0.728
2 - lr: 0.0010
Epoch 12/15
8/8 [=====] - 7s 884ms/step - loss: 0.5613 - accuracy: 0.6904 - val_loss: 0.5519 - val_accuracy: 0.631
1 - lr: 0.0010
Epoch 13/15
8/8 [=====] - 6s 784ms/step - loss: 0.5378 - accuracy: 0.7322 - val_loss: 0.5398 - val_accuracy: 0.660
2 - lr: 0.0010
Epoch 14/15
8/8 [=====] - 6s 766ms/step - loss: 0.5462 - accuracy: 0.6987 - val_loss: 0.5355 - val_accuracy: 0.718

```

FIGURE 11 : LES EPOCHS

Lors de l'entraînement du modèle nous remarquons que lorsque le nombre des époques est augmenté, le taux de précision augmente et le taux d'erreur diminue. Nous remarquons aussi que ceci n'est pas proportionnel car en arrivant à un certain seuil d'époques, le modèle commence à se stabiliser et l'augmentation du nombre d'époques n'est pas aussi importante qu'au début. Cependant nous avons remarqué également que l'augmentation des époques permet d'améliorer la classification des catégories mais le temps d'exécution augmente à chaque fois.

La matrice de confusion donne plus de détails sur les résultats de classification obtenus sur la base de données. La diagonale de cette matrice affiche le nombre d'images correctement classées et elle implique qu'elle a été bien construite par notre modèle. Cette matrice est obtenue en saisissant le code ci-dessous.

```
# Prédire Les classes de l'ensemble de validation
y_pred = classifieur.predict(valid_set)
y_pred = (y_pred > 0.5) # Convertit Les probabilités en classes binaires

# Obtenir Les vraies classes
y_true = valid_set.classes

# Créer La matrice de confusion
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_true, y_pred[:len(y_true)])

# Visualiser La matrice de confusion
sns.heatmap(cm, annot=True, cmap='Blues', fmt='g')
plt.xlabel('Predicted labels')
plt.ylabel('True labels')
plt.show()
```

4/4 [=====] - 1s 266ms/step

FIGURE 12 : CODE POUR LA MATRICE DE CONFUSION

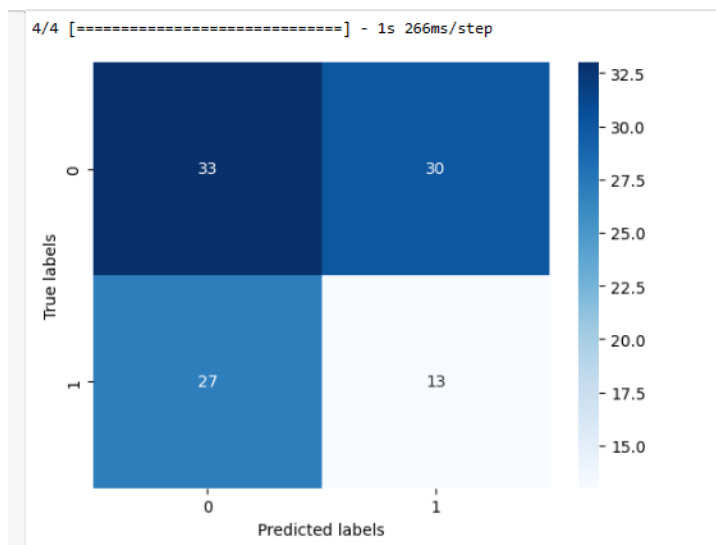


FIGURE 13 : MATRICE DE CONFUSION

```

Entrée [26]: import numpy as np
import keras.utils as image
test_image=image.load_img("test/malignant (31).png",target_size=(64,64))
test_image=image.img_to_array(test_image)
test_image=np.expand_dims(test_image,axis=0)
result=classifier.predict(test_image)
training_set.class_indices

1/1 [=====] - 0s 59ms/step

Out[26]: {'malignant': 0, 'normal': 1}

Entrée [27]: if result[0][0] == 0:
prediction = "malignant"
else:
prediction = "normal"

Entrée [28]: prediction

Out[28]: 'malignant'

```

FIGURE 14 : FONCTION DE PREDICTION

V. Faisabilité

L'étude de faisabilité est très importante du fait qu'elle aide le sponsor du projet, visant à contrôler si le projet est réalisable ou non, au travers de l'analyse de différents facteurs : Facteurs économiques et financiers. La budgétisation du projet passe par la budgétisation des différentes phases et aussi par la budgétisation de toutes leurs activités. Tout ceci ne concerne que le plan technique du projet

1. Technique

L'évaluation du coût du projet sera divisée entre les différentes ressources nécessaires qui sont : logicielles, matériels et humaines.

Logiciel	Rôle	Prix en Fcfa
Power BI	C'est un tableau pour la visualisation des données	30000
Windows	Système d'exploitation de Microsoft	125 00
Total		155000

TABLEAU 1 : RESSOURCES LOGICIELLE

La configuration du matériel a utilisé est la suivante :

Logiciel	Prix (FCFA)
PC (CPU 2.53GHz, 12Go RAM, 500Go ROM)	600 000 Une connexion
Internet	60 000 / mois
Total	700 000

TABLEAU 2: RESSOURCES MATERIELLES

La réalisation de ce projet nécessite la main d'œuvre d'un Data Analyst, d'après le site web Glassoor le salaire moyen pour le poste de Data Analyst est de 200 000 FCFA en Cameroun.

Conclusion

Dans ce chapitre, nous avons décrit la mise en place de la classification à l'aide du Deep Learning, centré sur le cancer du sein. Ce modèle consacre à la détection précoce de cette maladie en exploitant une base de données d 1600 images mammographiques. Nos tests ont souligné l'importance de la taille de la base de données pour la formation efficace des réseaux de neurones convolutifs. Ainsi, pour optimiser un modèle de Deep Learning, il est essentiel de disposer d'une vaste base de données pour obtenir des performances optimales.

CONCLUSION GENERALE

Rendus au terme de notre rapport, nous avons pu constater l'immense potentiel de l'IA dans le domaine médical. Dans le cadre de ma formation académique, nous avons eu l'opportunité unique d'intégrer l'équipe de Inoni Tech pour un stage. Cette expérience a été d'une grande valeur, non seulement pour appliquer les connaissances théoriques acquises pendant mes études, mais aussi pour comprendre les défis et les dynamiques du monde professionnel.

L'objectif central de notre travail était de concevoir un modèle sophistiqué capable de classer avec précision les images mammographiques, un outil crucial pour le diagnostic précoce du cancer du sein. En plongeant dans le domaine du traitement d'images, nous avons été impressionnés par les capacités du Deep Learning, en particulier compte tenu des avancées récentes dans ce domaine. Parmi les nombreuses techniques disponibles, nous avons choisi d'utiliser les réseaux de neurones convolutifs (CNN) pour notre modèle. Cette décision était basée sur leur simplicité relative, couplée à une capacité éprouvée à traiter et classer des images avec une précision exceptionnelle. Le résultat a été stupéfiant : notre modèle a affiché une précision de 71%, dépassant largement nos attentes initiales.

Au terme de nos investigations, nous sommes fiers de déclarer que non seulement nos suppositions étaient exactes, mais les objectifs que nous nous étions fixés pour cette recherche ont été atteints. Les implications de cette étude pour le futur de la détection du cancer du sein sont immenses et prometteuses.

REFERENCES WEBOGRAPHIQUES

- ✓ L'Organisation Mondiale de la Santé fournit des informations détaillées sur le cancer du sein, notamment les statistiques mondiales, les symptômes, les facteurs de risque et les traitements <https://www.who.int/fr/news-room/fact-sheets/detail/breast-cancer>
- ✓ Un article d'ONU Info rapporte que le cancer du sein est désormais la pathologie oncologique la plus fréquente dans le monde <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein>
- ✓ Comprendre le cancer du sein <https://ww5.komen.org/BreastCancer/UnderstandingBreastCancer.html>
- ✓ Breast Cancer Information and Resources <https://www.breastcancer.org/>
- ✓ Breast Cancer in Women <https://www.hopkinsmedicine.org/health/conditions-and-diseases/breast-cancer/breast-cancer-in-women>
- ✓ About Breast Cancer <https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html>
- ✓ <https://www.cancerresearchuk.org/about-cancer/breast-cancer.html>

REFERENCES BIBLIOGRAPHIQUES

- Puyuelo Laurent. « Bibliographie », dans : Cancer du sein. Un médecin à l'épreuve de l'annonce, sous la direction de Puyuelo Laurent. Toulouse, Érès, « L'Ailleurs du corps », 2011, p. [191-192](#)
- Francequin Ginette. « Bibliographie », dans : Cancer du sein : Une féminité à reconstruire. Sous la direction de Francequin Ginette. Érès, 2012, pp. [231-240](#)
- [Cairn.info fournit une liste de publications sur le cancer du sein, notamment des ouvrages et des articles scientifiques](#)

TABLE DE MATIERES

RESUME	2
ABSTRACT	3
SOMMAIRE	4
ACRONYMES ET SIGLES	6
LISTES DES FIGURES	7
LISTES DES TABLEAUX	7
INTRODUCTION GENERALE	8
I. Contexte général de l'étude	8
II. Problématique de l'étude	9
a. Présentation du problème	9
b. Problème général	9
c. Problèmes spécifiques	9
III. Hypothèse de l'étude	9
a. Hypothèse générale	9
b. Hypothèses spécifiques	9
IV. Objectif de l'étude	10
a. Objectif général l'étude	10
b. Objectif Spécifique de l'étude	10
c. Justification de l'Etude	10
d. Délimitation de l'étude	10
e. Plan du mémoire	11
CHAPITRE I : CADRE THEORIQUE ET CONCEPTUEL	12
Introduction	13
I. Cadre théorique et état d'art	13
II. Historique	14
III. Cadre Règlementaire	16
Conclusion	17
CHAPITRE II : Méthodologie de recherche	18
Introduction	19
I. Nature de recherche	19
II. Variable de recherche	19
a. Définition conceptuelle	19
Instruments de recherche	22

Conclusion.....	24
CHAPITRE III : PRESENTATION D'INONI TECH ET DES DONNEES COLLECTEES.....	25
Introduction	26
I. Présentation de l'entreprise.....	26
a. Produits et services de l'entreprise	26
b. Localisation de l'entreprise	26
c. Mission.....	27
d. Organigramme de l'entreprise.....	28
II. Données collectées.....	28
Conclusion.....	29
CHAPITRE IV : ANALYSE DIAGNOSTIQUE DE LA SITUATION ET PROPOSITION D'INTERVENTION	30
Introduction	31
I. Présentation et analyse de la situation.....	31
II. Intervention proposée et justification.....	32
III. Objectifs de l'intervention - projet envisagé	33
a. Objectif général	33
b. Objectifs spécifiques.....	33
IV. Composante de l'intervention envisagée	33
a. Stratégies d'action et contenu : Conception d'un modèle Deep Learning pour la détection du Cancer du sein.....	33
b. Résultats obtenus et discussion.....	35
V. Faisabilité	38
1. Technique.....	38
Conclusion.....	39
CONCLUSION GENERALE.....	40
REFERENCES WEBOGRAPHIQUES.....	41
REFERENCES BIBLIOGRAPHIQUES	41
TABLE DE MATIERES.....	42