

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et
Populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la
Recherche scientifique

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA 1

كلية التكنولوجيا
Faculté de Technologie

قسم إلكترونيك
Département d'Électronique



Ref : RT23

MEMOIRE DE MASTER

Filière : Télécommunication

Spécialité : Réseaux et Télécommunications

Présenté par :

BENSAFI Lyes Adel

Segmentation et classification des nodules thyroïdiens à partir des images échographiques en utilisant Deep Learning

Proposé par :

Dr. YAHIAOUI Amina

Année Universitaire : 2024-2025

Remerciements

*Après avoir rendu grâce à **Dieu** le Tout Puissant et le Miséricordieux nous tenons à remercier vivement tous ceux qui, de près ou de loin ont participé à la rédaction de ce document.*

*Je tiens à exprimer ma profonde gratitude à **Docteur Yahiaoui**, ma promotrice, pour s'accompagnement constant, ses conseils avisés et sa bienveillance tout au long de la réalisation de ce mémoire. Je souhaite également la remercier pour la qualité de son enseignement et l'engagement dont elle a toujours fait preuve durant mon parcours universitaire.*

*Mes remerciements les plus sincères vont également à **Madame Boutaleb** et **Docteur Berkat**, membres du jury, pour l'honneur qu'elles m'ont fait en acceptant d'évaluer ce travail. Leur présence est d'autant plus précieuse qu'elles ont également été mes enseignantes, et qu'elles ont grandement contribué à ma formation par la rigueur et la clarté de leurs cours.*

Enfin, j'adresse mes remerciements chaleureux à mes camarades de spécialité, ainsi qu'à tous les membres du département d'électronique, pour leur accompagnement, leur disponibilité et leur contribution à ma formation.

Dédicace

C'est avec une profonde gratitude et sincères mots que je dédie ce travail en premier lieu à Mes chères parents qui quel que soit l'objet qu'on essayera de leur offrir, il n'atteindra jamais ce que j'ai envie de leur dire et exprimer

A mon cher père ABDELGHANI

Celui qui est toujours là pour moi, et m'a donné un magnifique modèle de labeur et de persévérance, Celui qui m'a encouragé qui m'a toujours protégé, mon modèle qui fait ma fierté.

Que ce travail soit l'expression des vœux que tu n'as cessé de formuler dans tes prières. Mon cher père.

A ma chère mère ZAHIA

La perle de mon existence, quelle brave dame que tu es, que de sacrifices consentis à mon égard afin que je progresse dans mes études. Je tombe en admiration devant la bonté de ton cœur à nulle pareille.

Quels que soient mes caprices et mes écarts tu m'as toujours soutenue, trouvant les mots justes pour me ramener sur le bon chemin.

L'évènement que nous célébrons aujourd'hui t'est entièrement dédié. Que dieu te préserve santé et longue vie

*À mes frères **NABIL** et **HIND**, pour leur affection, leur compréhension et leur patience. Votre présence à mes côtés a été une source de motivation et de réconfort.*

ملخص

تمثل عقيدات الغدة الدرقية تحديًا طبيًا كبيرًا بسبب تواترها واحتمالية إصابتها بالأورام الخبيثة. تسلط هذه الأطروحة الضوء على إمكانات الذكاء الاصطناعي للمساعدة في تشخيص وإدارة المرضى الذين يعانون من هذا المرض، من خلال تطوير نظام ذكي يعتمد على التعلم العميق، قادر على تقسيم العقيدات تلقائيًا من صور الموجات فوق الصوتية، من أجل تصنيفها على أنها حميدة/خبيثة. تم تنفيذ بنى مختلفة YOLOv5 وDeepLabv3+ وAttention U-Net، وتم تقييمها على مجموعة بيانات TN3K. كان أداء نموذج U-Net هو الأفضل، حيث بلغ معامل النرد أكثر من 90% ودقة بنسبة 96%.

الكلمات المفتاحية: عقيدات الغدة الدرقية، التعلم العميق، التقسيم، التصنيف.

Résumé

Les nodules thyroïdiens représentent un enjeu médical important en raison de leur fréquence et de leur potentiel de malignité. Ce mémoire met en évidence le potentiel de l'intelligence artificielle dans l'aide au diagnostic et la prise en charge des patients atteints de cette maladie, en développant un système intelligent basé sur le Deep Learning, capable de segmenter automatiquement les nodules à partir des images échographiques, afin de les classer, par la suite, en bénin/malin. Différentes architectures ont été implémentées, à savoir U-Net, YOLOv5, DeepLabv3+, Attention U-Net, et évaluées sur le dataset TN3K. Le modèle U-Net s'est révélé le plus performant avec un coefficient Dice dépassant les 90% et une accuracy de 96 %.

Mots clés : Nodule thyroïdien, Deep Learning, Segmentation, Classification.

Abstract

Thyroid nodules represent a major medical challenge due to their frequency and potential for malignancy. This thesis highlights the potential of artificial intelligence to help diagnose and manage patients suffering from this disease, by developing an intelligent system based on Deep Learning, capable of automatically segmenting nodules from ultrasound images, in order to subsequently classify them as benign/malignant. Different architectures were implemented, namely U-Net, YOLOv5, DeepLabv3+, Attention U-Net, and evaluated on the TN3K dataset. The U-Net model performed best, with a Dice coefficient of over 90% and an accuracy of 96%.

Key words: Thyroid nodule, Deep Learning, Segmentation, Classification.

Table des matières

Liste des figures	ix
Liste des Tableaux	xi
Liste des Acronymes et Abréviations	xiii
Introduction générale.....	1
Chapitre I . Fondements Médicaux	4
I.1. Introduction	4
I.2. La glande thyroïdienne	4
I.2.1. Anatomie de la glande thyroïdienne	4
I.2.2. Rôle hormonal de la thyroïde.....	5
I.2.3. Pathologies fréquentes de la glande thyroïde	7
I.2.3.1. Hypothyroïdie	7
I.2.3.2. Hyperthyroïdie	8
I.2.3.3. Nodules thyroïdiens.....	8
I.3. Les nodules thyroïdiens.....	8
I.3.1. Types de nodules thyroïdiens	9
I.3.1.1. Classification morphologique échographique	9
I.3.1.2. Classification fonctionnelle	9
I.3.1.3. Classification histopathologie	10
I.3.2. Risques de malignité des nodules thyroïdiens	10
I.3.3. Données épidémiologiques sur les nodules thyroïdiens	11
I.4. Méthodes diagnostiques et sémilogie des nodules thyroïdiens par imagerie médicale	12
I.4.1. Diagnostic par Échographie Ultrasonore (US)	12
I.4.1.1. Principes techniques de l'échographie thyroïdienne :	12
I.4.2. Diagnostic par Tomodensitométrie (Scanner).....	13
I.4.2.1. Indications spécifiques du scanner thyroïdien :	14
I.4.3. Diagnostic par Imagerie par Résonance Magnétique (IRM)	15
I.5. Conclusion.....	16
Chapitre II . L'Apport du Deep Learning dans la segmentation et la classification des nodules thyroïdiens	18
II.1 Introduction	18
II.2 Notion d'image	18
II.2.1 Définition de l'image numérique	19

II.2.2	Les caractéristiques d'une image numérique	19
II.2.2.1	Pixel.....	19
II.2.2.2	Dimension	20
II.2.3	Les Types d'image	21
II.2.3.1	Image binaire	21
II.2.3.2	Image en niveaux de gris	21
II.2.3.3	Image en couleurs (RGB)	22
II.3	Relation IA, ML et DL.....	22
II.4	Deep Learning en imagerie médicale.....	24
II.5	Réseaux de neurones :.....	25
II.5.1	Réseaux neuronaux récurrent	25
II.6	Techniques de segmentation	27
II.6.1	Méthodes traditionnelles.....	27
II.6.2	Le modèle U-Net.....	28
II.7	Modelés de Deep Learning pour la classification	Erreur ! Signet non défini.
II.7.1	Classification binaire	29
II.7.2	Réseaux neuronaux convolutifs	29
II.8	Métrique d'évaluation	30
II.9	Conclusion.....	31
Chapitre III	Environnement de développement et modèle d'apprentissage	33
III.1	Introduction.....	33
III.2	Environnement de développement.....	33
III.2.1	Matériel utilisé.....	33
III.2.2	Environnement de simulation	34
III.3	Base de données.....	37
III.3.1	Présentation générale du jeu de données TN3K	37
III.3.2	Structure du jeu de données	37
III.4	Prétraitement des images	38
III.5	Modèle de Segmentation	40
III.5.1	Architecture générale d'U-Net	40
III.5.2	Caractéristiques clés.....	40
III.5.3	Justification du choix	41
III.5.4	Comparaison avec d'autres architectures.....	41
III.6	Fonction d'activation.....	42

III.7	Fonction de perte	43
III.8	Modèle de Classification	44
III.8.1	Architecture du modèle de classification	44
III.9	Conclusion	47
Chapitre IV . Résultats expérimentaux et évaluation du modèle		49
IV.1	Introduction.....	49
IV.2	Choix de jeu de données	49
IV.2.1	Avantages du jeu de données TN3K :	50
IV.3	Analyse comparative des modèles et justification du choix final pour la segmentation.....	51
IV.3.1	Expérimentation 1 « 1 ^{ère} évaluation avec le modèle DeepLabv3+ ».....	51
IV.3.2	Expérimentation 2 « 2 ^{ème} évaluation avec le modèle Yolov5+ XGBoost »	54
IV.3.3	Expérimentation 3 « 3 ^{ème} évaluation avec le modèle Attention U-Net ».....	56
IV.3.4	Expérimentation 4 « Évaluation du modèle U-Net avec techniques d'amélioration »..	58
IV.4	Étapes de réalisation du modèle de segmentation final.....	59
IV.5	Paramètres du modèle d'apprentissage profond	62
IV.6	Paramètres d'architecture du modèle U-Net final.....	63
IV.7	Présentation des résultats du modèle final	64
IV.8	Comparaison des modèles	66
IV.9	Jeu de données AUITD.....	68
IV.9.1	Introduction.....	68
IV.9.2	Limites du dataset AUITD pour une classification approfondie	68
IV.9	CNN pour la classification binaire des nodules thyroïdiens (malin / bénin)	69
IV.9.3	Performance élevée sur les images médicales.....	69
IV.9.4	Moins de prétraitement requis	70
IV.9.5	Bonne capacité de généralisation	70
IV.9.6	Validation par la littérature scientifique	70
IV.10	Étapes de réalisation du modèle de classification	70
IV.11	Paramètres du modèle CNN utilisée	73
IV.12	Paramètres d'architecture du modèle CNN	75
IV.13	Analyse des performances du modèle de classification.....	76
IV.13.1	Courbes d'apprentissage.....	76
IV.13.2	Matrice de confusion.....	77
IV.13.3	Visualisation des erreurs	77
IV.13.4	Discussion et limitations.....	78

IV.14	Conception d'une Application Web	79
IV.14.1	Remarques.....	81
IV.15	Limites de la classification dans notre cas spécifique	82
IV.16	Conclusion	82
Conclusion Générale.....		85
Perspectives.....		86
Bibliographie.....		88

Liste des figures

Chapitre I : fondement médicaux

Figure I-1 : Illustration anatomique et schématique de la glande thyroïde.

Figure I-2 : Structure microscopique de la glande thyroïde.

Figure I-3 : Physiologie thyroïdienne.

Figure I-4 : Classification fonctionnelle des nodules thyroïdiens.

Figure I-5 : Schéma récapitulatif des principaux facteurs de risque de malignité des nodules thyroïdiens.

Figure I-6 : Photos échographique des nodules thyroïdiens.

Figure I-7 : Photos de scanner pour une glande thyroïde.

Figure I-8 : Photo IRM pour une glande thyroïde.

Chapitre II : L'Apport du Deep Learning dans la segmentation et la classification des nodules Thyroïdiens

FigureII-1 : Image numérique.

Figure II -2 : Pixels d'une image en niveaux de gris.

Figure II -3 : Dimension d'une image numérique.

Figure II -4 : Image Binaire.

Figure II -5 : Image en niveau de gris.

Figure II -6 : Image couleur avec une représentation matricielle RVB.

Figure II -7 : Relation entre AI, ML, et Deep Learning.

Figure II -8 : Extraction des caractéristiques DL vs ML.

Figure II -9 : Schéma d'un neurone Biologique à gauche et représentation d'un neurone artificiel à droit.

Figure II-10 : Architecture d'un RNN.

Figure II-11 : Auto-encoder Architecture.

Figure II-12 : Architecture U-NET.

Figure II-13 : Schéma représentant l'architecture d'un CNN.

Chapitre III : Environnement de développement et modèle d'apprentissage

Figure III-1 : Effet de l'égalisation d'histogramme sur une image échographique.

Figure III-2 : Architecture d'un CNN.

Figure III-3 : couche de convolution.

Figure III-4 : Représentation schématique de l'opération de pooling dans un CNN.

Chapitre IV : Résultats expérimentaux et évaluation du modèle

Figure IV-1 : Courbes d'apprentissage DeepLabv3+.

Figure IV-2 : résultats de segmentation du modèle DeepLabv3+.

Figure IV-3 : Résultat de simulation de modèle YOLOv5+GXboots sur kaggle.

Figure IV-4 : résultats de segmentation du modèle Attention U-Net.

Figure IV-5 : Courbes d'apprentissage Attention U-Net.

Figure IV-6 : Précision du modèle Attention U-Net.

Figure IV-7 : Courbes d'apprentissage du modèle U-Net.

Figure IV-8 : Organigramme méthodologique du modèle U-Net.

Figure IV-9 : résultats de segmentation du modèle final.

Figure IV-10 : Courbes d'apprentissage du modèle final.

Figure IV-11 : Précision du modèle final.

Figure IV-12 : Organigramme méthodologique du modèle CNN.

Figure IV-13 : Courbes d'apprentissage du modèle CNN.

Figure IV-14 : Matrice de confusion pour le modèle CNN.

Figure IV-15 : résultats de classification du modèle.

Figure IV-16 : L'interface d'application.

Figure IV-17 : Résultat de segmentation d'une image du nodule via l'interface.

Figure IV-18 : Résultat de classification d'une image du nodule via l'interface.

Figure IV-19 : Résultat de segmentation + classification d'une image du nodule via l'interface.

Liste des Tableaux

Chapitre I : fondement médicaux

Tableau I-1 : Fonctions physiologique des hormones thyroïdiennes.

Tableau I-2 : Classification morphologique échographique des nodules thyroïdiens.

Chapitre II : L'Apport du Deep Learning dans la segmentation et la classification des nodules Thyroïdiens

Tableau II-1 : les principales métriques utilisées.

Chapitre III : Environnement de développement et modèle d'apprentissage

Tableau III-1 : Spécifications techniques du matériel utilisé.

Tableau III-2 : Avantages de l'utilisation de la plateforme Kaggle.

Tableau III-3 : Critères et Justifications du Choix de Streamlit.

Tableau III-4 : Raisons de non-utilisation d'Anaconda.

Tableau III-5 : Caractéristiques principales du dataset TN3K.

Tableau III-6 : Tableau récapitulatif des étapes de prétraitement.

Tableau III-7 : Caractéristiques principales d'U-Net.

Tableau III-8 : Comparaison des architectures basées sur U-Net pour la segmentation des nodules thyroïdiens.

Tableau III-9 : structure du modèle CNN.

Chapitre IV : Résultats expérimentaux et évaluation du modèle

Tableau IV-1 : Comparaison des jeux de données pour la segmentation des nodules.

Tableau IV-2 : Paramètres d'entrainement du modèle DeepLabv3+.

Tableau IV-3 : Paramètres d'entrainement du modèle Yolov5+XGboost.

Tableau IV-4 : Paramètres d'entrainement du modèle Attention U-Net.

Tableau IV-5 : Paramètres des données et prétraitement du modèle final.

Tableau IV-6 : Paramètres d'architecture du modèle final.

Tableau IV-7 : Comparaison synthétique des modèles testés.

Tableau IV-8 : Paramètres d'entraînement du modèle de classification par CNN.

Tableau IV-9 : Paramètres d'architecture du modèle CNN.

Liste des Acronymes et Abréviations

AI	I ntelligence A rtificielle.
ANN	A rtificiel N eural N etworks.
AUC	A rea U nder the C urve.
AUITD	A lgerian U ltrasound I mages T hyroid D ataset.
BCE	B inary C ross- E ntropy.
CNN	R éseaux de N eurons C onvolution.
DL	D eep L earning.
ECG	É lectro C ardio G ramme.
FNA	F ine N eedle A spiration.
GRU	G âte R écurrent U nit.
GPU	U nité de T raitement G raphique.
IRM	I magerie par R ésonance M agnétique
IoU	I ntersection O ver U nion.
LSTM	L ong S hort- T erme M emory.
ML	M achine L earning.
Png	P ortable N etwork G raphics.
RVB	R ouge, V ert, B leu.
RGB	R ed, G reen, B lue.
RNN	R écurrent N eural N etworks
TSH	T hyroïde S timulation H ormone.
TI-RADS	T hyroid I maging R eporting A nd D ata S ystem.
TN3K	T hyroid N odule 3000.
US	U ltrasons.
VAE	V ariational A utoencoders.
VRAM	M émoire V idéo à A ccès A léatoire.

INTRODUCTION GENERALE

Introduction générale

La glande thyroïde joue un rôle essentiel dans le maintien de l'équilibre physiologique de l'organisme, et les anomalies qui l'affectent, en particulier les nodules thyroïdiens, soulèvent des enjeux médicaux importants. Bien que majoritairement bénins, certains nodules peuvent présenter un risque de malignité, rendant leur détection et leur caractérisation précises cruciales pour une prise en charge optimale.

Dans ce contexte, les avancées récentes en intelligence artificielle, notamment dans le domaine de l'apprentissage profond (*Deep Learning, DL*), offrent des outils puissants pour automatiser l'analyse des images échographiques. Ces technologies permettent non seulement d'améliorer la sensibilité du diagnostic, mais aussi de réduire les erreurs humaines et les actes invasifs inutiles.

L'objectif de ce mémoire est de développer un système intelligent basé sur des modèles de Deep Learning, capable de réaliser la segmentation et la classification automatique des nodules thyroïdiens à partir d'images échographiques. Ce travail s'inscrit dans une démarche de recherche appliquée visant à contribuer à l'amélioration des outils d'aide au diagnostic dans un cadre médical concret.

Le manuscrit est structuré en quatre chapitres :

Le premier chapitre présente les fondements médicaux indispensables à la compréhension des pathologies thyroïdiennes, en particulier les nodules, ainsi que les méthodes actuelles de diagnostic.

Le deuxième chapitre introduit les concepts clés de l'intelligence artificielle, du Deep Learning et de leur application à l'imagerie médicale, avec un focus sur les architectures adaptées à la segmentation et à la classification d'images.

Le troisième chapitre décrit l'environnement de développement, les outils logiciels, les bibliothèques utilisées ainsi que la méthodologie adoptée pour l'implémentation des modèles.

Le quatrième chapitre est consacré à l'évaluation expérimentale. Il présente les résultats obtenus lors de l'entraînement de plusieurs modèles (U-Net, DeepLabv3+, YOLOv5, etc.), ainsi qu'une analyse comparative des performances.

Enfin, une conclusion générale synthétise les apports de ce travail, met en lumière ses limites et propose des perspectives pour de futures améliorations ou recherches.

Chapitre I.

Fondements Médicaux

Chapitre I . Fondements Médicaux

I.1. Introduction

Le système endocrinien joue un rôle fondamental dans la régulation des fonctions vitales de l'organisme. Parmi les glandes qui le constituent, la thyroïde occupe une position centrale en raison de ses fonctions métaboliques, développementales et homéostatiques [1].

Ce chapitre vise à présenter les fondements médicaux nécessaires à la compréhension de cette glande, la physiologie et les pathologies thyroïdiennes, notamment les nodules, afin d'éclairer les approches actuelles d'évaluation et de prise en charge.

I.2. La glande thyroïdienne

La thyroïde, en forme de papillon, est une glande endocrine située à la base du cou, en avant de la trachée. Grâce à la sécrétion d'hormones thyroïdiennes, elle joue un rôle crucial dans la régulation du métabolisme, de la croissance, du développement du système nerveux et de nombreuses autres fonctions physiologiques. Une compréhension approfondie de cette glande passe par l'étude de son anatomie et de ses sécrétions hormonales [2].

I.2.1. Anatomie de la glande thyroïdienne

La thyroïde est caractérisée par sa forme rappelant un papillon avec deux lobes latéraux (lobe droit et lobe gauche) réunis par une partie centrale appelée l'isthme, ainsi illustrée dans la figure (I-1). Son poids moyen chez l'adulte est d'environ 15 à 30 grammes, et ses dimensions moyennes sont d'environ 4 cm de large sur 2 cm de haut pour chaque lobe [2].

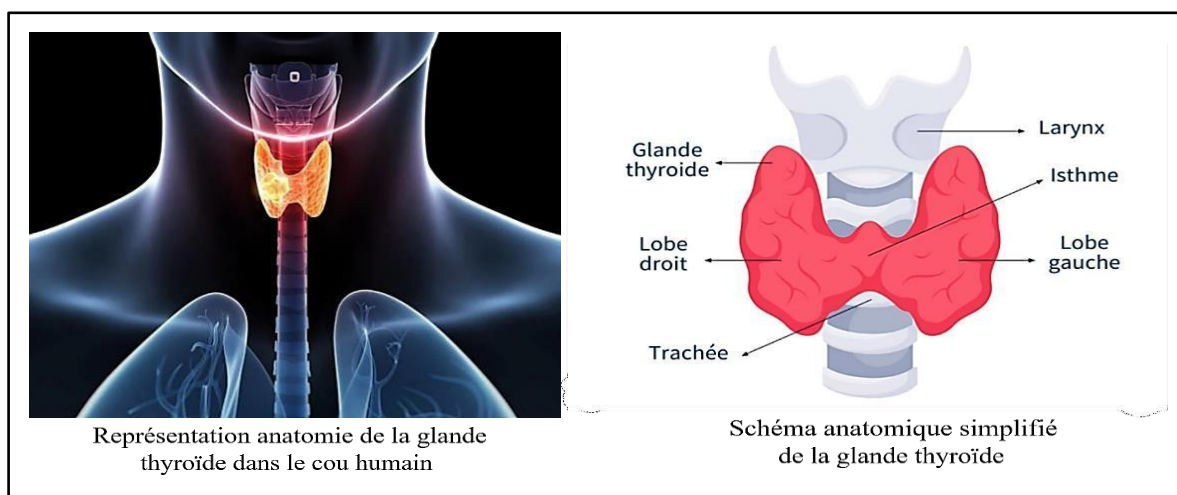
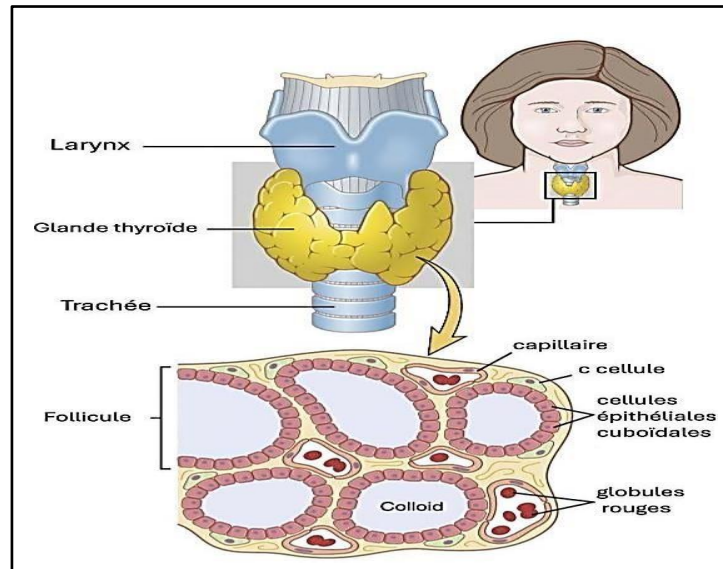


Figure I-1 : Illustration anatomique et schématique de la glande thyroïde.

Microscopiquement, la thyroïde est composée principalement de structures sphériques appelées follicules thyroïdiens, figure (I-2). Chaque follicule contient une substance gélatineuse appelée colloïde, entourée d'une seule couche de cellules épithéliales cuboïdales, responsables de la synthèse et de la sécrétion des hormones thyroïdiennes (thyroxine T4 et



triiodothyronine T3).

Figure I-2 : Structure microscopique de la glande thyroïde [1].

Entre ces follicules, on retrouve les cellules C (cellules parafolliculaires), qui sécrètent la calcitonine, hormone jouant un rôle essentiel dans la régulation du taux sanguin de calcium. La vascularisation abondante représentée par les nombreux capillaires et les globules rouges permet une libération efficace des hormones dans la circulation sanguine [3].

I.2.2. Rôle hormonal de la thyroïde

Les hormones thyroïdiennes, principalement la thyroxine (T4) et la triiodothyronine (T3), jouent des rôles fondamentaux dans la régulation des fonctions physiologiques de l'organisme [3]. Ainsi présentée dans la figure (I-3), les hormones thyroïdiennes (T3 et T4) jouent un rôle clé dans le maintien de l'équilibre physiologique global du corps humain. Toute perturbation de leur activité peut rapidement altérer le fonctionnement général de l'organisme, soulignant ainsi l'importance d'une régulation précise et continue [3].

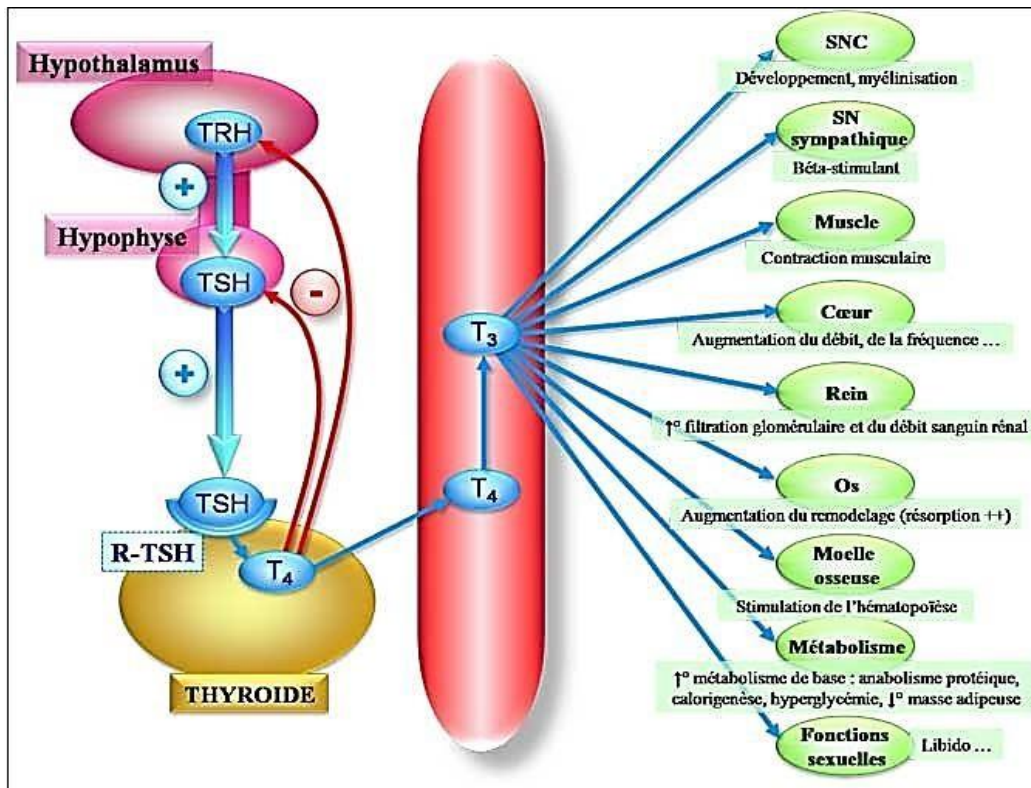


Figure I-3 : Physiologie thyroïdienne.

Le tableau (I-1) explique les principales fonctions physiologiques des hormones thyroïdiennes :

Fonction physiologique	Effets
Régulation du métabolisme basal et énergétique	- Augmentation du métabolisme basal et de la consommation d'oxygène - Stimulation du catabolisme des glucides, lipides et protéines - Maintien de la thermogénèse corporelle (production de chaleur)
Développement neurologique et croissance	- Développement du système nerveux central chez le fœtus et le nourrisson - Croissance osseuse et musculaire durant l'enfance et l'adolescence

Régulation cardiovasculaire	- Accélération du rythme cardiaque et augmentation du débit cardiaque - Dilatation des vaisseaux sanguins, réduisant ainsi la résistance vasculaire périphérique
Régulation du système digestif	- Stimulation de la motilité gastro-intestinale, facilitant la digestion et l'absorption
Influence sur le système nerveux central et l'humeur	- Régulation de l'humeur, de l'attention et des fonctions cognitives - Effet sur le cycle veille-sommeil et l'état émotionnel
Autres fonctions physiologiques	- Régulation de la reproduction et de la fertilité - Contrôle de la maturation et différenciation cellulaire - Régulation de l'équilibre hydrique et électrolytique

Tableau I-1 : Fonctions physiologique des hormones thyroïdiennes.

I.2.3. Pathologies fréquentes de la glande thyroïde

La glande thyroïde, bien qu'étant de petite taille, joue un rôle central dans la régulation de nombreuses fonctions physiologiques. Elle peut être le siège de plusieurs pathologies fréquentes, parmi lesquelles l'hypothyroïdie, l'hyperthyroïdie et les nodules thyroïdiens. Ces dysfonctionnements sont à l'origine de déséquilibres hormonaux parfois sévères, ayant des impacts significatifs sur la qualité de vie et la santé générale [4].

I.2.3.1. Hypothyroïdie

L'hypothyroïdie désigne une insuffisance de production des hormones thyroïdiennes (principalement T3 et T4), ce qui entraîne une baisse généralisée du métabolisme basal. Elle est l'une des pathologies thyroïdiennes les plus courantes, particulièrement chez les femmes et les personnes âgées. Les causes fréquentes incluent la thyroïdite auto-immune de Hashimoto, la chirurgie thyroïdienne, l'irradiation cervicale, et une carence en iode. Cliniquement, les patients présentent une fatigue chronique, une frilosité, une constipation, une prise de poids,

une peau sèche, un ralentissement psychomoteur et des troubles menstruels. Sur le plan biologique, l'hypothyroïdie se caractérise par une augmentation de la TSH (thyroid-stimulating hormone) et une diminution des taux de T3 et T4 [5].

I.2.3.2. Hyperthyroïdie

L'hyperthyroïdie est une condition opposée, où la glande thyroïde produit une quantité excessive d'hormones thyroïdiennes, entraînant une accélération du métabolisme. Elle est fréquemment causée par la maladie de Basedow (forme auto-immune), un goitre multinodulaire toxique ou un adénome toxique. Les signes cliniques sont caractéristiques : perte de poids malgré un appétit accru, tachycardie, nervosité, insomnie, tremblements, intolérance à la chaleur, diarrhée, aménorrhée, et parfois exophtalmie dans le cadre de la maladie de Basedow. Le diagnostic repose sur un dosage hormonal (TSH basse, T3 et T4 élevées) et des examens d'imagerie pour identifier l'origine de la suractivité [6].

I.2.3.3. Nodules thyroïdiens

Un nodule est une boule, plus ou moins volumineuse, située au sein de la thyroïde. Il peut être constitué par du tissu thyroïdien et/ou du liquide. Lorsqu'il contient du liquide, c'est alors un kyste. Il peut y avoir un seul, plusieurs ou de nombreux nodules (goitre multinodulaire). En règle générale, les nodules sont de petite taille, indolores, perturbant rarement le fonctionnement de la thyroïde et n'entraînant pas de signe d'hyperthyroïdie. Le goitre multinodulaire, quant à lui, peut comprimer les organes avoisinants s'il grossit trop. S'il mesure plus de 1 cm de diamètre, le nodule peut être palpé ; au-dessous de cette taille, il peut uniquement être détecté à l'échographie [3].

Les nodules sont le plus souvent asymptomatiques. Cependant certains peuvent provoquer une gêne, une dysphagie, voire une compression locale. L'évaluation se fait généralement par échographie selon le système TI-RADS, et la cytoponction (FNA) est recommandée pour les nodules suspects. Les progrès récents en intelligence artificielle visent à optimiser la caractérisation automatique de ces lésions à partir d'images échographiques, réduisant les biopsies inutiles et améliorant la détection des cas malins [4].

I.3. Les nodules thyroïdiens

Cette section explore les bases de la classification morphologique et fonctionnelle des nodules, les systèmes de risque comme le TI-RADS, ainsi que les outils actuels permettant d'affiner le diagnostic et d'orienter les décisions thérapeutiques.

I.3.1. Types de nodules thyroïdiens

Les nodules thyroïdiens sont des formations de taille variable, de composition variable, présentes dans le tissu thyroïdien. Ils peuvent être classés selon leur fonction biologique, leur contenu, voire leurs caractéristiques histopathologiques. Cette classification est nécessaire pour évaluer le risque de cancer et orienter la prise de décision [4].

I.3.1.1. Classification morphologique échographique

Du fait qu'elle est non invasive, l'échographie est l'examen référent pour l'examen des nodules thyroïdiens. Elle permet de caractériser avec précision leur structure, leur composition ainsi leurs contours. Ces éléments sont utilisés pour classer les nodules morphologiquement en trois grands types : kystiques, solides ou mixtes (solido-kystiques). Cette classification est importante car elle permet d'estimer le risque de malignité et d'organiser la stratégie d'exploration, notamment la réalisation d'une cytoponction [7].

I.3.1.2. Classification fonctionnelle

En plus de leur morphologie, les nodules de la glande thyroïde peuvent être caractérisés par leur activité fonctionnelle, fonctionnelle que l'on détermine soit par un examen clinique, soit par la scintigraphie à l'iode radioactif. Le nodule dit chaud sécrète des hormones en plus grande quantité que la normale, tandis que, dans le nodule froid, elles sont soit en quantité normale, soit moins sécrétées. C'est l'un des critères de la malignité, mais aussi de l'orientation diagnostique. L'organigramme de la figure (I-4) ci-après, illustre cette dichotomie fonctionnelle [8] [9] :

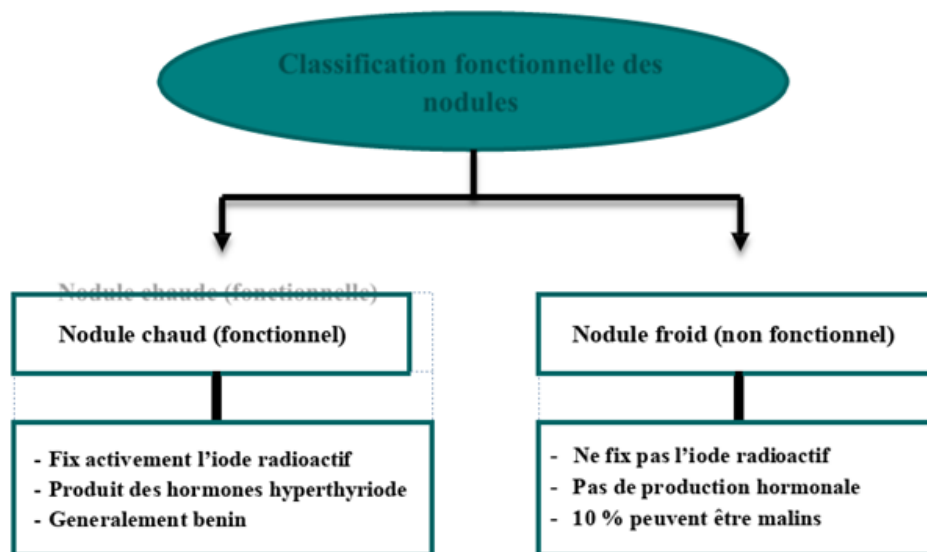


Figure I-4 : Classification fonctionnelle des nodules thyroïdiens.

I.3.1.3. Classification histopathologie

L'analyse histopathologique repose sur l'examen microscopique des cellules constituant les nodules thyroïdiens. Elle permet la distinction de deux grandes catégories : les nodules bénins, de loin les plus fréquents, et généralement sans gravité, et les nodules malins qui reflètent différents types de cancers thyroïdiens [10].

✚ Nodules bénins :

Faisant partie des nodules colloïdaux, des adénomes folliculaires, et de l'hyperplasie nodulaire notamment. Ils sont les plus fréquents, ne nécessitant qu'une surveillance dans la grande majorité des cas.

✚ Nodules malins :

Nodules présentant plusieurs types de cancers de la thyroïde :

- Carcinome papillaire (qui représente le plus fréquent (~85 % des cancers thyroïdiens)).
- Carcinome folliculaire.
- Carcinome médullaire (moins souvent observé (cancer d'origine parafolliculaire)).
- Carcinome anaplasie (rare et très agressif).

I.3.2. Risques de malignité des nodules thyroïdiens

Les nodules thyroïdiens sont des lésions courantes mais sont le plus souvent de nature bénigne, néanmoins 5 à 15 % d'entre eux sont malins, d'où la nécessité d'une bonne évaluation pour le diagnostic précoce de ces cancers, surtout du carcinome papillaire qui représente près

de 85 % des cancers thyroïdiens différenciés [8]. Le risque de malignité est fonction de plusieurs éléments cliniques et radiologiques sont presentes dans la figure (I-5) :

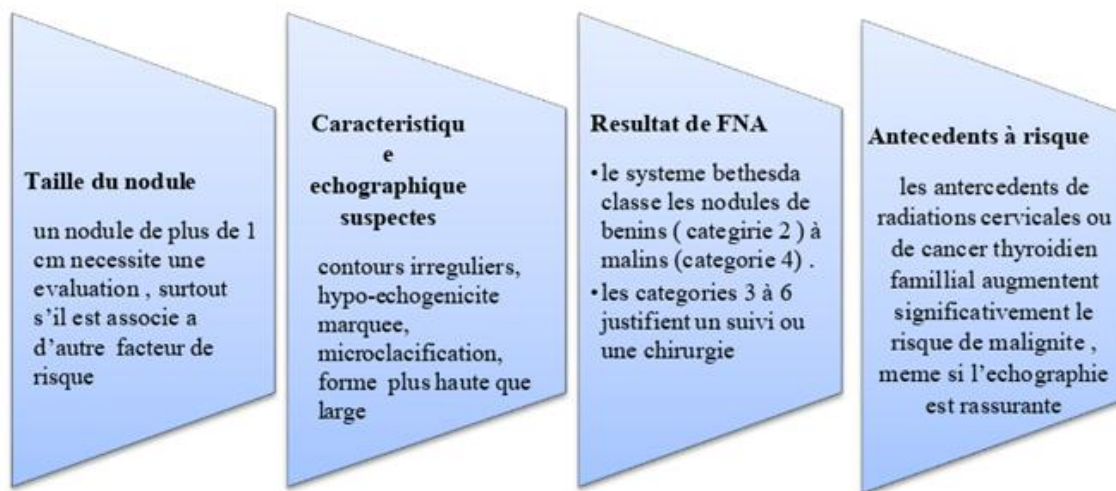


Figure I-5 : Schéma récapitulatif des principaux facteurs de risque de malignité des nodules thyroïdiens.

I.3.3. Données épidémiologiques sur les nodules thyroïdiens

Au sein de la population générale, les nodules thyroïdiens représentent une affection fréquente. Les estimations les plus largement communiquées évaluent entre 20 et 70 % les adultes ayant au moins un nodule thyroïdien décelable à l'échographie, mais la majorité des nodules sont retrouvés fortuitement, asymptomatiques [4] [6]. La prévalence des nodules augmente de façon significative avec l'âge, le sexe féminin, les antécédents d'irradiation cervicale ainsi que la carence en iode.

La plupart des nodules sont acquises et bénins, mais entre 5 et 15 % des nodules présentent un risque potentiel de malignité, justifiant la mise en place de protocoles stricts de dépistage et de stratification du risque. Par ailleurs, des données épidémiologiques récentes montrent une augmentation apparente de l'incidence des cancers thyroïdiens dans plusieurs pays, en partie expliquée par l'amélioration des techniques d'imageries (échographie haute résolution), mais aussi par le phénomène de surdiagnostic du fait de la détection de microcarcinomes qui n'auraient probablement jamais été cliniquement significatifs [8].

Les résultats des travaux montrent que la fréquence des nodules thyroïdiens est deux à quatre fois plus élevée en ce qui concerne la gent féminine que la gent masculine, atteignant

50 % pour les patientes âgées de plus de 60 ans. Une telle constatation justifie également la pratique d'une échographie systématique, normalisée pour éviter des gestes invasifs non justifiés, ainsi qu'une détection précoce des lésions malignes à propos desquelles [9].

I.4. Méthodes diagnostiques et séméiologie des nodules thyroïdiens par imagerie médicale

La très forte prévalence des nodules thyroïdiens parmi la population générale l'impose un diagnostic rigoureux afin de bien identifier les lésions bénignes des lésions potentiellement néoplasiques. Actuellement, les techniques d'imagerie médicales occupent une place essentielle dans l'évaluation initiale, la bonne caractérisation et dans le suivi des nodules thyroïdiens. L'échographie (US), la tomodensitométrie (scanner) et l'imagerie par résonance magnétique (IRM) constituent les principales techniques utilisées. Chacune a des avantages spécifiques et est indiquée dans des contextes cliniques précis pour permettre un diagnostic précis et un traitement adapté [4].

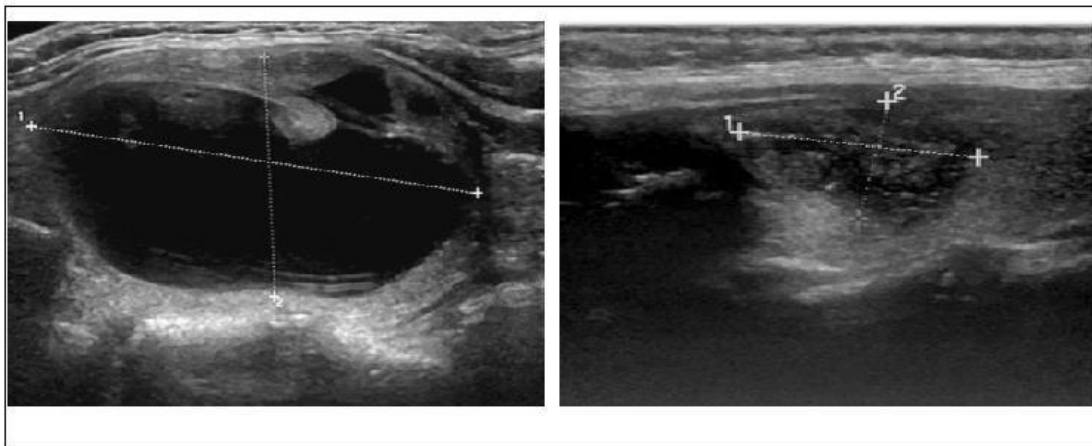


Figure I-6 : Images échographique des nodules thyroïdiens .

I.4.1. Diagnostic par Échographie Ultrasonore (US)

L'échographie est la modalité diagnostique de première intention pour l'évaluation initiale des nodules thyroïdiens, grâce à sa facilité d'accès, son innocuité (absence de radiation), son faible coût et son excellente résolution des tissus superficiels [7].

I.4.1.1. Principes techniques de l'échographie thyroïdienne :

L'échographie utilise des ultrasons émis par une sonde placée sur la peau du cou du patient pour générer des images détaillées de la thyroïde et identifier précisément les caractéristiques internes des nodules :

Échogénicité :

Les nodules hypoéchogènes présentent un risque accru de malignité par rapport aux nodules hyperéchogènes ou isoéchogènes [7].

Forme du nodule « taller-than-wide » :

Les nodules présentant une hauteur supérieure à leur largeur sont fortement associés au cancer thyroïdien [5].

Contours et marges :

Les marges irrégulières, lobulées ou spiculées augmentent le risque de malignité [4].

Présence de calcifications :

La présence de microcalcifications est une caractéristique fortement suspecte de malignité, souvent associée au carcinome papillaire [7].

Vascularisation par Doppler :

L'utilisation du Doppler couleur aide à identifier une hypervascularisation intranodulaire, généralement associée à une suspicion plus élevée de cancer.

L'avènement de l'intelligence artificielle et des réseaux neuronaux convolutifs appliqués à l'échographie permet une analyse automatisée et plus objective, réduisant les erreurs liées à l'interprétation humaine et diminuant les biopsies inutiles [5].

I.4.2. Diagnostic par Tomodensitométrie (Scanner)

La tomodensitométrie (scanner) utilise des rayons X pour produire des images précises en coupes axiales permettant une excellente évaluation anatomique des structures adjacentes à la thyroïde. Malgré son irradiation ionisante, le scanner est particulièrement utile dans des situations spécifiques [5].



Figure I-7 : Photos de scanner pour une glande thyroïde

I.4.2.1. Indications spécifiques du scanner thyroïdien :

✚ Évaluation anatomique approfondie :

Utile pour les nodules volumineux plongeants ou rétrosternaux qui ne peuvent être entièrement visualisés par échographie.

✚ Détection d'envahissement local :

Le scanner permet de définir avec précision l'invasion éventuelle des structures adjacentes (trachée, œsophage, vaisseaux du cou), crucial pour la planification chirurgicale.

✚ Évaluation préopératoire et post-opératoire :

Détection des métastases ganglionnaires cervicales et médiastinales, utiles pour la stadification et le suivi du cancer thyroïdien avancé.

Les limites du scanner restent principalement l'exposition aux radiations ionisantes et la nécessité fréquente de produits de contraste iodés (risque allergique et thyrotoxicose induite chez certains patients [5]).

I.4.3. Diagnostic par Imagerie par Résonance Magnétique (IRM)

L'IRM utilise un champ magnétique puissant et des ondes radio pour générer des images à fort contraste tissulaire. Elle est particulièrement bénéfique dans des situations où l'échographie ou le scanner sont limités, en raison de son excellente résolution anatomique et absence d'irradiation ionisante [11].



Figure I-8 : Photo IRM pour une glande thyroïde .

✚ Évaluation des nodules complexes ou invasifs :

IRM utile pour les nodules infiltrant les tissus mous cervicaux ou en cas de suspicion d'invasion des structures vasculaires ou neurologiques adjacentes.

✚ Alternative aux patients sensibles ou contre-indiqués au scanner :

Patients allergiques aux produits iodés, insuffisants rénaux, femmes enceintes ou jeunes patients sensibles aux radiations.

✚ Évaluation des ganglions cervicaux profonds :

Identification précise des ganglions envahis par les cancers thyroïdiens invasifs.

Les limitations de l'IRM sont principalement son coût élevé, sa durée d'acquisition prolongée et son accessibilité limitée [11].

I.5. Conclusion

Le contenu de ce chapitre a permis d'expliquer l'anatomie la physiologie et les principales pathologies de la glande thyroïdienne, notamment les nodules thyroïdiens, fréquents mais parfois source d'inquiétude en raison du risque de malignité potentielle. La connaissance de ces éléments et du vocabulaire médico-chirurgical est une condition nécessaire pour comprendre les outils de diagnostic utilisés, notamment l'échographie, le scanner, l'IRM, mais aussi en raison des progrès technologiques de l'intelligence artificielle qui permettent de mieux évaluer certaines lésions thyroïdiennes de façon plus précise et personnalisée. Cette transition naturelle vers les méthodes d'imageries sera traitée dans le chapitre suivant.

CHAPITRE II.

L'Apport du Deep Learning dans la segmentation et la classification des nodules thyroïdiens

Chapitre II . L'Apport du Deep Learning dans la segmentation et la classification des nodules thyroïdiens

II.1 Introduction

L'intelligence artificielle (AI), et plus particulièrement le Deep Learning (DL), connaît un essor spectaculaire dans le domaine de sante, en particulier pour le traitement et l'analyse des images médicales. Grâce à sa capacité à apprendre directement à partir de données brutes, le Deep Learning s'impose aujourd'hui comme un outil incontournable dans l'aide au diagnostic, notamment en échographie. De plus, le dépistage et la classification des nodules thyroïdiens sont cruciaux pour écarter les cancers thyroïdiens, dont l'incidence continue d'augmenter dans le monde.

En dépit des outils diagnostiques actuels, le challenge demeure de distinguer avec précision entre les lésions bénignes et malignes, et c'est là que les avancées récentes en DL, s'avèrent prometteuses. L'intégration de l'analyse d'images par des modèles d'apprentissage profond pourrait améliorer significativement la précision du diagnostic et réduire le besoin de biopsies invasives.

Ce chapitre vise à explorer les fondements theoriques du Deep Learning, en mettant particulièrement l'accent sur ses applications dans le domaine de l'imagerie medicale. Dans un premier temps, un lexique des principaux termes lies au traitement d'image sera introduit afin de se familiariser avec le vocabulaire technique de base. Ensuite, nous passerons en revue les principales architectures de reseaux de neurones frequemment utilises dans la litterature. Nous analyserons egalement les differentes approches d'applications de ces modeles en imagerie medicale.

II.2 Notion d'image

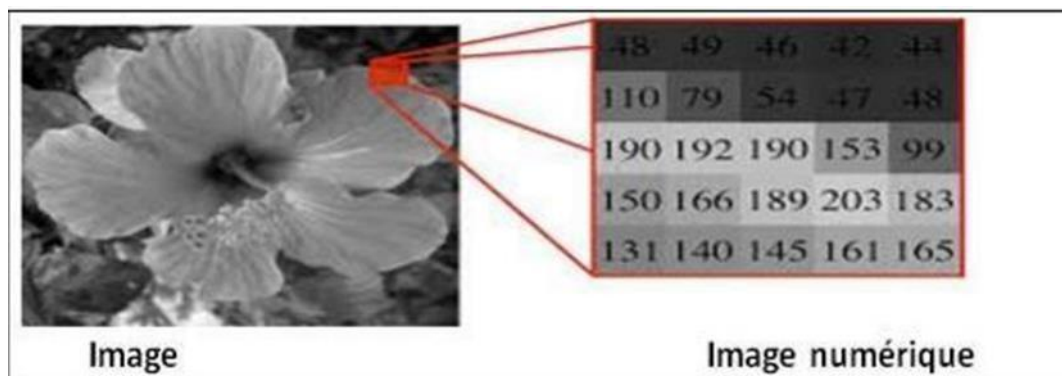
Une image peut être définie comme une représentation visuelle d'un objet, d'une scene ou d'une structure, capturée par des moyens physiques ou électroniques tels que la photographie, l'image médicale, le dessin vidéo. D'un point de vue mathématique, une image est souvent modélisée comme une fonction bidimensionnelle continue $I(x, y)$, où x et y représentent les coordonnées spatiales, et I est une fonction d'intensité lumineuse (et parfois de couleur).

Toutefois, pour être traitée numériquement par une machine, cette représentation doit être discrétisée, c'est-à-dire convertie en une matrice de pixels à valeurs numériques par un processus de numérisation [12].

II.2.1 Définition de l'image numérique

Une image numérique est une représentation d'une scène visuelle, modélisée comme une fonction à support discret (souvent bidimensionnel) et à valeurs discrètes. Elle est généralement structurée sous forme de matrice de pixels, où chaque pixel représente une unité de surface avec une intensité codée par une valeur entière, ainsi illustrée dans la figure (II.1).

Les images peuvent être scalaires, comme les images en niveaux de gris, ou vectorielles, comme les images couleurs (Rouge, Vert et Bleu RVB) ou multi spectrales. Lorsqu'on effectue un zoom sur une image numérique, on observe clairement la structure discrète en pixels, soulignant son caractère numérique par opposition aux images analogiques [12].



FigureII-1 : Image numérique

II.2.2 Les caractéristiques d'une image numérique

Une image numérique est définie par un ensemble de caractéristiques fondamentales telles que le pixel, qui en constitue l'unité de base, et la dimension, qui détermine sa taille et sa résolution spatiale.

II.2.2.1 Pixel

Le terme pixel, contraction de (Picture élément), désigne la plus petite unité constitutive d'une image numérique. Chaque pixel correspond à un point de l'image et contient une valeur numérique représentant une intensité lumineuse (pour les images en niveaux de gris) ou un vecteur de composantes (pour les images couleur). La qualité d'une image dépend en grande

partie du nombre de pixels (résolution), influençant ainsi sa finesse et son niveau de détail. La figure (II.2) ci-dessous montre un exemple de pixel d'une image en niveaux de gris [12].

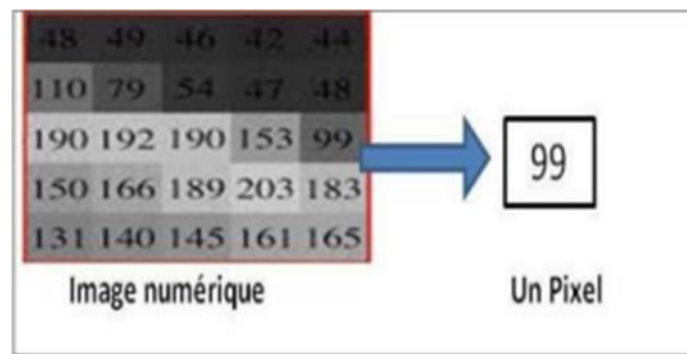


Figure II -2 : Pixels d'une image en niveaux de gris

II.2.2.2 Dimension

La dimension d'une image numérique fait référence à la taille de sa grille de pixels, généralement exprimée en termes de largeur $\times$ hauteur. Par exemple, une image de dimension 640×480 contient 640 pixels sur l'axe horizontal (colonnes) et 480 pixels sur l'axe vertical (lignes) soit un total de 307200 pixels. Cette dimension détermine la résolution spatiale de l'image. Elle influence directement la quantité d'information visuelle que peut contenir une image. La figure (II.3) montre un exemple minimaliste du sens de dimensionnement d'une image [12].

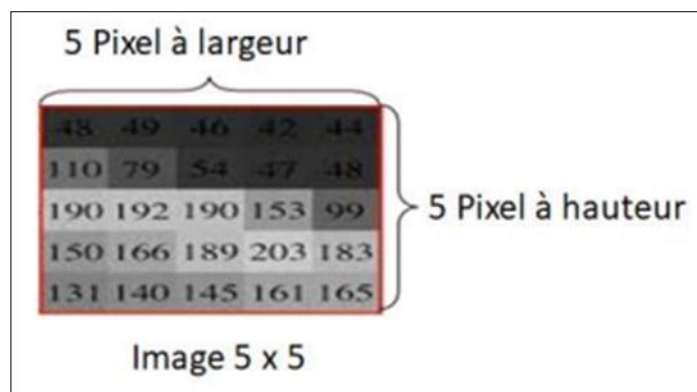


Figure II -3 : Pixels d'une image numérique

II.2.3 Les Types d'image

Les images numériques se déclinent en plusieurs types selon la nature et le codage de l'information qu'elles contiennent. Les plus courantes sont les images binaires, les images en niveaux de gris et les images en couleurs (RVB), chacune répondant à des besoins spécifiques en traitement d'image [12].

II.2.3.1 Image binaire

Une image binaire est un type d'image numérique dans laquelle chaque pixel ne peut prendre que deux valeurs 0 et 1. Généralement, le 0 représente le noir et le 1 représente le blanc, figure (II-4). Ce type d'image est couramment utilisé dans les tâches de segmentation, de masquage, ou d'analyse de formes, ou l'on souhaite simplement distinguer la présence ou d'un objet.

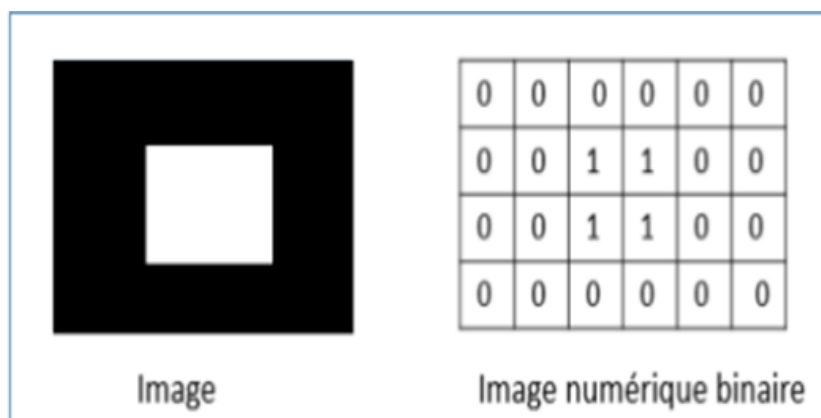


Figure II -4 : Image Binaire.

II.2.3.2 Image en niveaux de gris

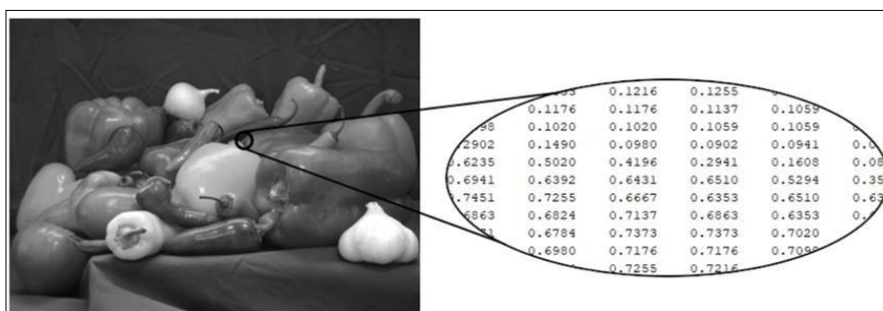


Figure II -5 : Image en niveau de gris.

C'est une image dans laquelle chaque pixel représente une intensité lumineuse unique, sans information de couleur. Les valeurs des pixels varient généralement entre 0 (correspondant au noir) et 255 (correspondant au blanc), avec un total de 256 valeurs, montrées dans la figure (II-5). Ce type d'image est largement utilisé en traitement des images médicales, car il permet de représenter efficacement les structures anatomiques à contraste variable tout en réduisant la complexité des données.

II.2.3.3 Image en couleurs (RGB)

Une image couleur est composée de pixels dont chacun est défini par un triple de composantes : Rouge(R), Vert (V), Bleu(B). Ce système appelé RVB (ou bien RGB en anglais), repose sur le mélange additif de ces trois couleurs primaires pour produire une large gamme de couleurs. Chaque composante est généralement codée sur 8 bits, prenant ainsi une valeur entière entre 0 et 255, ainsi illustrée dans la figure (II-6). Ce modèle est largement utilisé en informatique graphique, en photographie numérique et en traitement d'image couleur.



Figure II -6 : Image couleur avec une représentation matricielle RVB [Site officiel Matlab].

II.3 Relation IA, ML et DL

Apparue pour la première fois en 1955, l'intelligence artificielle (IA) est une vaste branche de l'informatique qui s'intéresse à la construction de machines/outils intelligents capables d'effectuer des tâches qui requièrent généralement l'intervention de l'homme [13].

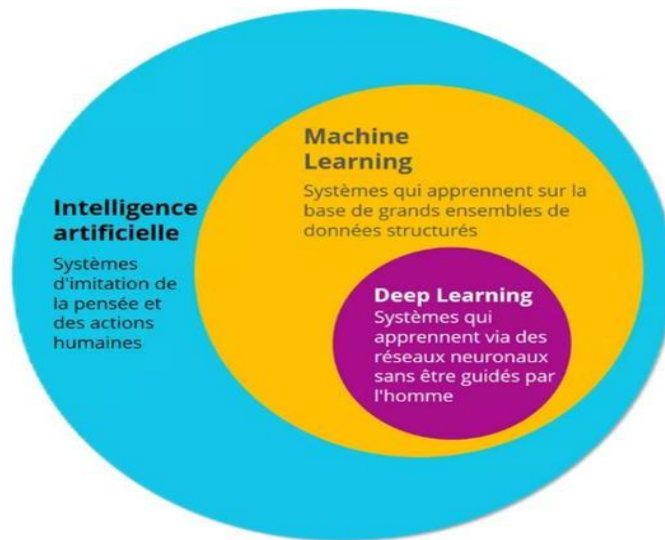


Figure II -7 : Relation entre AI, ML, et Deep Learning.

L'apprentissage automatique (machine Learning, en anglais, ML) est un sous-domaine de l'intelligence artificielle, figure (II.7), où des algorithmes permettent aux machines d'apprendre à partir de données sans être explicitement programmées. Il inclut :

- ***Apprentissage supervisé*** : Le modèle est entraîné avec des données étiquetées.
- ***Apprentissage non supervisé*** : Utilisé pour détecter des patterns dans des données non étiquetées.

L'apprentissage profond (Deep Learning, en anglais, DL) est une branche avancée de l'apprentissage automatique, repose sur des structures appelées réseaux de neurones artificiels. Les réseaux de neurones profonds, en particulier, se composent de multiples couches de neurones artificiels, chaque couche effectuant une transformation non linéaire de données d'entrée, ce qui leur confère la capacité d'extraction de caractéristiques à différents niveaux d'abstraction. La figure (II-8) présente la différence dans le traitement des données entre l'apprentissage automatique et l'apprentissage profond [14].

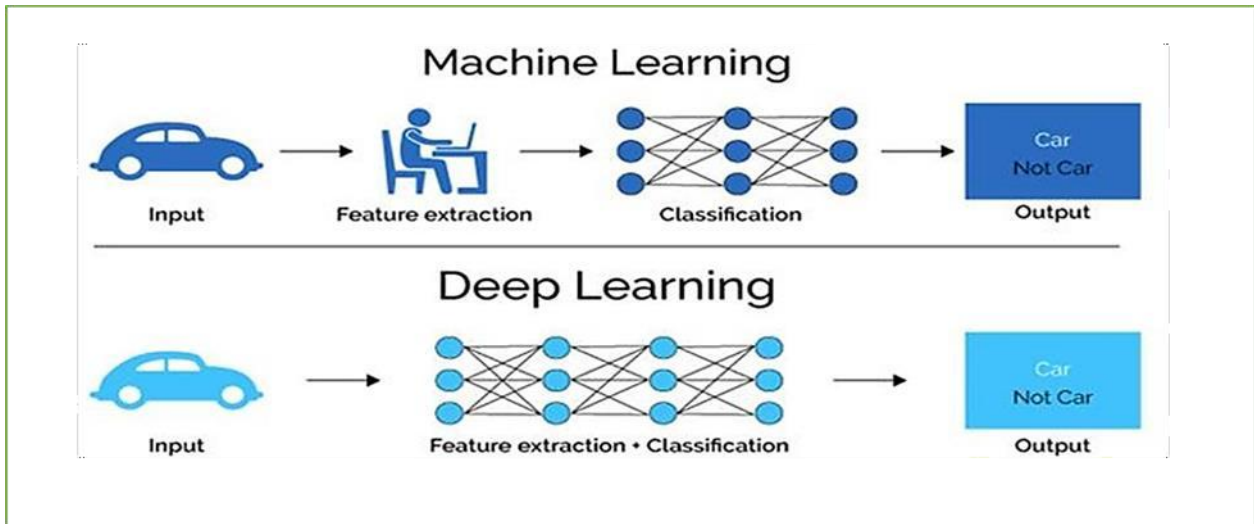


Figure II -8 : Extraction des caractéristiques DL vs ML

II.4 Deep Learning en imagerie médicale

Un aspect crucial du DL réside dans l'application de l'apprentissage supervisé, dont des modèles sont formés à partir d'exemples étiquetés. Cette procédure est essentielle dans le contexte médical, car elle influence la manière dont les algorithmes sont capables de généraliser et d'améliorer leurs performances en fonction des données cliniques disponibles, tout en minimisant les risques d'erreurs diagnostiques.

Les avancées récentes dans les algorithmes et l'augmentation de la puissance de calcul ont révolutionné l'application du deep learning dans le domaine médical, marquant une véritable rupture avec les méthodes traditionnelles de diagnostic, en offrant des outils capables de traiter des volumes de données non seulement variés, mais aussi en constante évolution [15].

En outre, des réseaux neuronaux profonds ont été utilisés pour la détection de cancers de la peau [15], la segmentation d'organes en IRM, ou encore pour la détection des nodules pulmonaires en radiographie. En ce qui concerne les nodules thyroïdiens, ces modèles peuvent être formés pour distinguer les nodules, leurs contours et ainsi distinguer leur nature à partir des échographies, permettant ainsi une prise de décision plus rapide et fiable pour les interventions cliniques [5].

Ces avancées ouvrent la voie à l'émergence d'outils décisionnels assistés par IA, contribuant à une médecine plus précise et personnalisée. L'intégration du DL dans ces systèmes de support à la décision clinique pourrait également contribuer à réduire la charge de travail des professionnels de santé tout en augmentant la qualité des soins fournis aux patients.

II.5 Réseaux de neurones :

Les réseaux de neurones artificiels (Artificial Neural Networks – ANN) sont des modèles d'apprentissage automatique inspirés du fonctionnement des neurones biologiques. Ils sont constitués de neurones artificiels (ou unités) organisés en couches (input, hidden, output), et interconnectés par des poids ajustables. Ces réseaux permettent d'apprendre des relations complexes entre des données d'entrée et des sorties cibles à travers un processus d'apprentissage supervisé ou non supervisé [14].

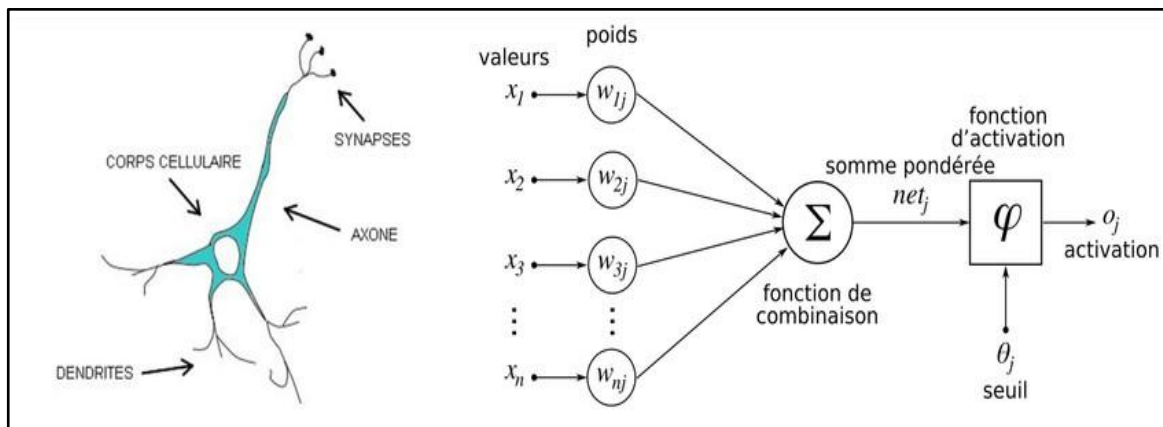


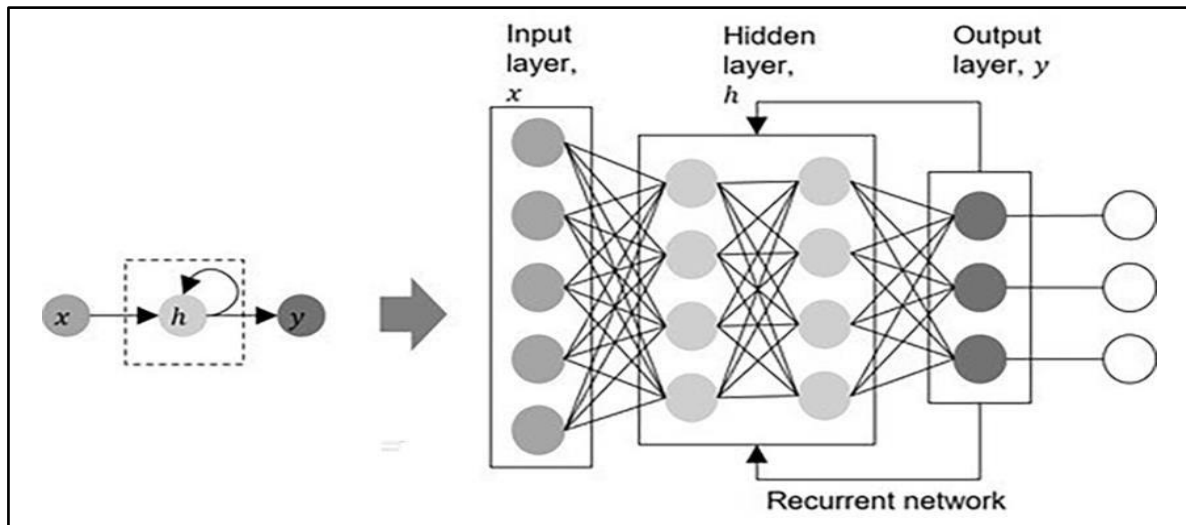
Figure II -9 : Schéma d'un neurone Biologique à gauche et représentation d'un neurone artificiel à droit

Selon leur structure et leur domaine d'application, plusieurs architectures de réseaux de neurones ont été développées. Voici les plus pertinents pour l'analyse d'image médicale, notamment dans le cadre des échographies de la thyroïde

II.5.1 Réseaux neuronaux récurrent

Les Réseaux neuronaux récurrent (RNN – Récurrent Neural Networks) sont conçus pour traiter des données séquentielles, c'est-à-dire des données qui dépendent du temps ou de l'ordre, comme des séries temporelles, des signaux ECG ou des séquences des images (vidéo médicale).

Comme le montre la figure (II-10), les RNN possèdent une mémoire interne, ce qui leur permet de prendre en compte le contexte précédent dans les prédictions. Ils sont souvent associés par des variantes comme LSTM (Long Short-term memory) et GRU (Gâte Récurrent



Unit), qui résolvent le problème du gradient qui disparaît, rendant l'apprentissage sur des longues séquences plus stable [16]. Bien que moins utilisées en imagerie médicale statique, les RNN peuvent intervenir dans les analyses de séquences échographique ou dans l'intégration du facteur temps pour suivre l'évolution d'un nodule sur plusieurs examens [16].

Figure II-10 : Architecture d'un RNN

II.5.2 Auto-encodeurs

Les auto-encodeurs sont des réseaux non supervisés capables de reconstruire leur entrée après un passage par une représentation compressée (encodage) et leur application en imagerie médicale peut être résumée en :

- compression des images médicales.
- Debruitage : suppression des artefacts et bruits sur les images d'échographie
- Prétraitement : extraction de caractéristiques non supervisée, pouvant être réutilisées dans des modèles de classification

Les variational autoencoders (VAE) constituent une version probabiliste très utilisée pour la génération des données médicales synthétique [17].

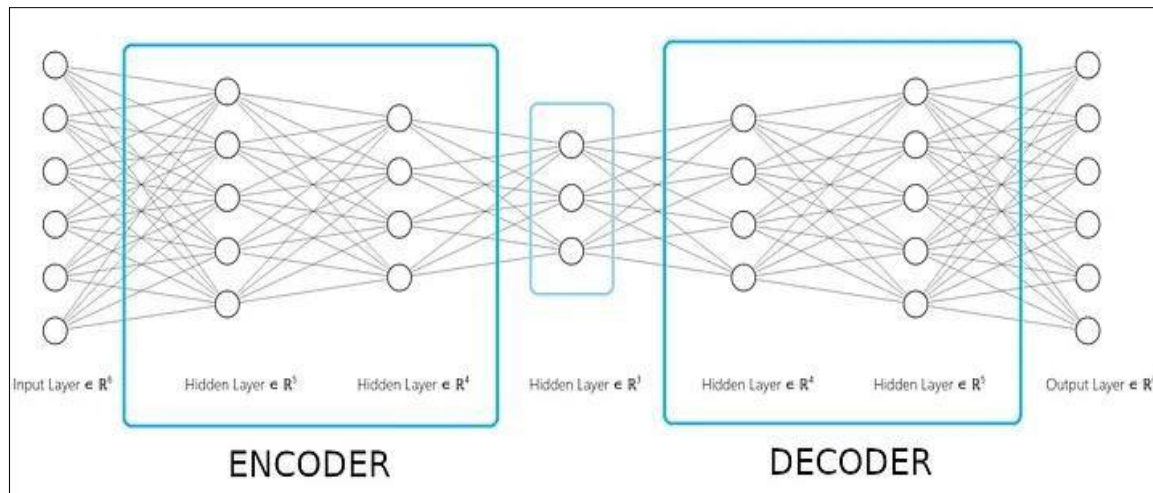


Figure II-11 : Auto-encodeur Architecture

II.6 Techniques de segmentation

La segmentation des images médicales est une étape cruciale dans le processus d'analyse et d'interprétation des données médicales, car elle permet d'isoler des structures ou des régions d'intérêt dans une image.

II.6.1 Méthodes traditionnelles

Parmi les techniques de segmentation, on distingue principalement la segmentation basée sur les pixels, qui consiste à classifier chaque pixel de l'image selon la classe à laquelle il appartient, souvent grâce à des méthodes de seuillage ou à des algorithmes comme K-means. Cette approche peut cependant se heurter à des problèmes de bruit ou de variations d'éclairage. La segmentation par contours, quant à elle, vise à détecter les bords des objets en utilisant des techniques de détection des contours, comme Canny et Sobel, permettant ainsi de délimiter les structures anatomiques. Cependant, cette méthode nécessite une bonne définition des contours, ce qui peut être très difficile dans des images présentant un faible contraste ou des artefacts.

Les dernières avancées en matière de segmentation reposent sur les techniques d'apprentissage profond, qui exploitent des architectures de réseaux de neurones convolutionnels (CNN) pour apprendre à segmenter des images à partir de grands ensembles de données annotées [18]. Parmi les architectures les plus couramment utilisées, on trouve le modèle U-Net, qui a été spécialement conçu pour la segmentation biomédicale. Cette méthode sera détaillée dans le titre suivant.

II.6.2 Le modèle U-Net

Le modèle U-Net, illustré dans la figure (II-12), a été spécifiquement conçu pour la segmentation d'images médicales, avec une architecture symétrique en forme de 'U'. Cette structure combine un chemin d'encodage (pour l'extraction de caractéristiques) et un chemin de décodage (pour la reconstitution de la segmentation), avec des connexions de saut (skip connections) qui facilitent la récupération des détails spatiaux fins. Ce modèle s'est révélé très efficace pour les tâches médicales nécessitant peu de données annotées comme en échographie thyroïdienne [19].

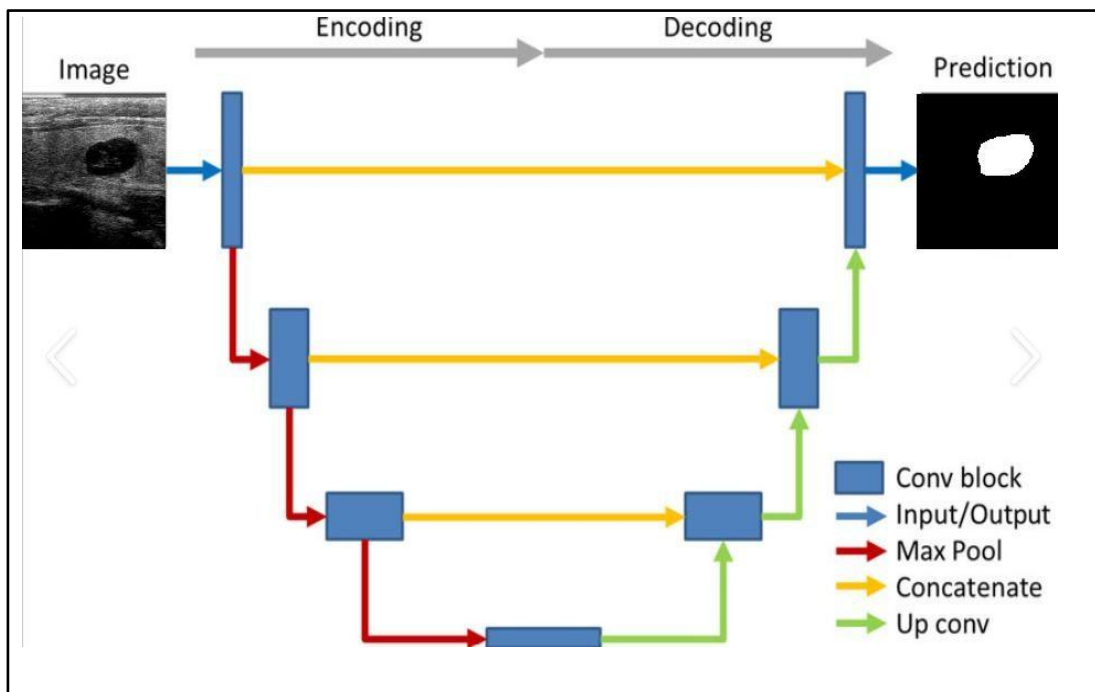


Figure II-12 : Architecture U-NET

II.7 Modelés de Deep Learning pour la classification

La classification des nodules thyroïdiens est une étape décisive pour distinguer les lésions bénignes des formes malignes. Grâce aux réseaux de neurones profonds, et en particulier aux CNN, cette tâche peut être automatisée avec une grande précision, réduisant ainsi les erreurs de diagnostic [5].

II.7.1 Classification binaire

Elle repose généralement sur une approche binaire, consistant à distinguer les nodules bénins (non cancéreux), malins (potentiellement cancéreux). Cette tâche est essentielle pour réduire les biopsies inutiles et guider les décisions cliniques. Dans ce contexte, les réseaux de neurones convolutés (CNN) tels que ResNet50, VGG16 ou encore DenseNet sont largement utilisés pour extraire automatiquement des caractéristiques discriminantes à partir des images échographiques [5].

Ces modèles profonds permettent une meilleure généralisation et surpassent les méthodes manuelles de détection fondées sur des critères morphologiques ou texturaux.

II.7.2 Réseaux neuronaux convolutifs

Les Réseaux neuronaux convolutés (CNN – Convolutional Neural Networks) sont aujourd'hui la référence pour le traitement d'images. Ils ont été introduits par LeCun (1998) avec le réseau LeNet, puis popularisés par AlexNet. Contrairement aux ANN classiques, les CNN utilisent des couches de convolution pour extraire automatiquement les caractéristiques spatiales locales (comme les bords, textures, formes) d'une image [20].

Son principe de fonctionnement est présenté dans la figure (II-9). Il se résume ainsi :

- Une couche de convolution applique des filtres sur l'image pour produire des cartes de caractéristiques.
- Des couches de pooling réduisent la taille de la carte tout en conservant l'information essentielle
- Des couches entièrement connectées en fin de réseau réalisent la classification finale
- Des fonctions d'activation comme ReLU introduisent la non-linéarité nécessaire à l'apprentissage de patterns complexes

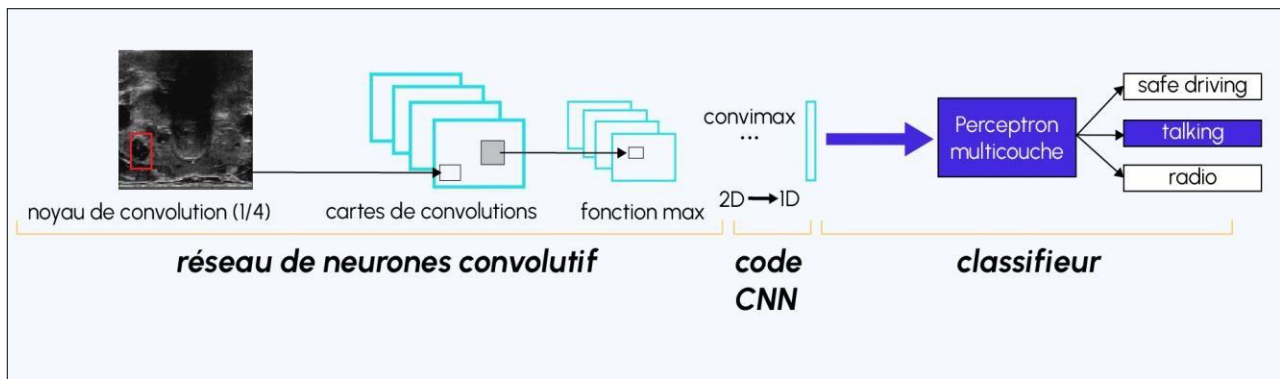


Figure II-13 : Schéma représentant l'architecture d'un CNN.

II.8 Métrique d'évaluation

L'évaluation de la performance d'un modèle de classification repose sur plusieurs indicateurs statiques. Le tableau ci-dessous résume les principales métriques utilisées, accompagnées de leurs formules mathématiques :

Métrique	Description	Formule
Accuracy	Pourcentage global de prédictions correctes.	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$
AUC (Area Under the Curve)	Aire sous la courbe ROC. Elle reflète la capacité du modèle à distinguer entre classes positives et négatives à différents seuils.	Aucune formule simple (calculée à partir de la courbe ROC)
Précision (Precision)	Proportion des prédictions positives qui sont réellement correctes.	$Precision = \frac{TP}{TP + FP}$
Rappel (Recall)	Proportion des cas réellement positifs correctement identifiés.	$Recall = \frac{TP}{TP + FN}$
F1-score	Moyenne harmonique entre précision et rappel. Utile en cas de classes déséquilibrées.	$F1 - score = 2 \times \frac{precision \times Recall}{precision + Recall}$

Tableau II-1 : les principales métriques utilisées

Avec : **TP** : vrais positifs ; **TN** : vrais négatifs ; **FP** : faux positifs ; **FN** : faux négatifs

Ces métriques permettent d'avoir une vision complète des performances du modèle, en tenant compte à la fois de la justesse des prédictions, de la capacité à identifier les cas positifs, et de robustesse en cas de déséquilibre des classes [14].

II .9 Conclusion

En conclusion la compréhension des différentes architecture des réseaux de neurones permet de choisir les modèles les mieux adaptés aux problématiques de segmentation et de classification des nodules thyroïdiens.

Ce chapitre a permis d'explorer les fondements théoriques et les applications pratiques du Deep Learning dans le domaine de l'imagerie médicale, avec un focus particulier sur les tâches de segmentation et de classification des nodules thyroïdiens. À Travers l'analyse des différentes architectures de réseaux de neurones profonds, il est apparu que le Deep Learning offre des solutions performantes et robustes pour automatiser ces processus critiques dans le diagnostic échographique.

CHAPITRE III.

Environnement de développement et modèle d'apprentissage

Chapitre III . Environnement de développement et modèle d'apprentissage

III.1 Introduction

Après avoir exposé les fondements théoriques du Deep Learning appliqué à la segmentation des images échographiques dans le chapitre précédent, il convient maintenant de détailler l'environnement technique et les outils utilisés pour la mise en œuvre pratique de notre système de segmentation des nodules thyroïdiens. Ce chapitre présente donc les choix effectués concernant le matériel informatique, l'environnement logiciel, la base de données utilisée ainsi que l'architecture du modèle d'apprentissage retenu.

III.2 Environnement de développement

Afin de mettre en œuvre le modèle de segmentation retenu, un environnement technique adapté a été configuré, combinant des ressources matérielles performantes et des outils logiciels spécialisés.

III.2.1 Matériel utilisé

Les expérimentations présentées dans ce travail de recherche ont été réalisées sur un ordinateur portable HP Victus 17, disposant des spécifications techniques citées dans le tableau (III-1) suivant :

<i>Composant</i>	<i>Spécification</i>
<i>Processeur</i>	AMD Ryzen 5 5600H (6 cœurs, 12 threads), cadencé à 3.3 GHz avec un boost jusqu'à 4.2 GHz.
<i>Carte graphique</i>	NVIDIA GeForce GTX 1650, équipée de 4 Go de mémoire vidéo GDDR6.
<i>Mémoire vive (RAM)</i>	16 Go DDR4, fonctionnant à une fréquence de 3200 MHz.
<i>Stockage</i>	Disque SSD de 1 To

Tableau III-1 : Spécifications techniques du matériel utilisé

Ce matériel offre une puissance de calcul suffisante pour effectuer des phases de développement locales, notamment le test initial du modèle U-Net sur des images échographiques de taille réduite à moyenne. Cependant, en raison des limites imposées par la capacité de la VRAM (notamment pour des batchs importants ou des résolutions élevées), une partie des entraînements a été réalisée sur la plateforme **cloud Kaggle**, qui fournit un accès gratuit à des ressources GPU plus performantes (NVIDIA Tesla T4).

III.2.2 Environnement de simulation

L'ensemble des simulations, entraînements et évaluations du modèle ont été réalisé sur la plateforme **Kaggle**, en utilisant **le langage Python** pour le développement et l'intégration du modèle, ainsi que **Streamlit** pour la création de l'application web destinée à la visualisation interactive des résultats obtenus.

III.2.2.1 Plateforme de développement : Kaggle

La plateforme Kaggle a été choisie comme environnement principal d'exécution pour plusieurs raisons pratiques et techniques :

<i>Avantages</i>	<i>Description</i>
<i>Accès gratuit aux GPU</i>	Permet d'accélérer considérablement l'entraînement des modèles CNN grâce à des cartes graphiques puissantes (NVIDIA Tesla T4 ou P100).
<i>Environnement Python préconfiguré</i>	Propose un espace de travail basé sur des notebooks Jupyter avec les bibliothèques principales déjà installées (TensorFlow, Keras, OpenCV, NumPy, Pandas, Matplotlib, etc.).
<i>Facilité de gestion des données</i>	Intégration native avec les jeux de données hébergés sur la même plateforme, permettant un chargement rapide et simple via des scripts Python.
<i>Reproductibilité et partage</i>	Les notebooks peuvent être facilement sauvegardés, versionnés et partagés, ce qui facilite la collaboration et la validation du code.

Tableau III-2 : Avantages de l'utilisation de la plateforme Kaggle

Grâce à ces avantages, Kaggle s'est révélé particulièrement adapté à notre projet, notamment pour pallier les limitations matérielles locales liées à la carte graphique (GTX 1650) et permettre un entraînement efficace du modèle U-Net sur des images échographiques.

III.2.2.2 Langage de programmation : Python

Le langage Python a été choisi comme langage principal de développement, en raison de ses nombreux avantages dans le domaine du deep learning et du traitement d'images médicales soulignant ainsi, en plus de grande communauté, sa simple syntaxe et ses bibliothèques spécialisées, notamment :

- **TensorFlow / Keras** : Pour la conception et l'entraînement du modèle U-Net.
- **OpenCV** : Pour le prétraitement des images (redimensionnement, normalisation, égalisation d'histogramme).
- **NumPy / Pandas** : Pour la manipulation efficace des matrices et des tableaux de données.
- **Matplotlib / Seaborn** : Pour la visualisation des courbes d'apprentissage et des résultats prédits.

III.2.2.3 Interface utilisateur : Application Web avec Streamlit

Afin de rendre les résultats de segmentation accessibles et interactifs, une application web a été développée à l'aide du Framework Streamlit pour plusieurs raisons, citées dans le tableau (III-3) ci-dessous :

<i>Critères</i>	<i>Justification</i>
<i>Simplicité d'utilisation</i>	Il ne nécessite aucune connaissance en HTML/CSS/JavaScript, ce qui est idéal pour un développeur débutant ou un chercheur souhaitant rapidement déployer une interface.
<i>Intégration naturelle avec Python</i>	L'application peut être construite à partir d'un script Python classique, sans configuration complexe.

<i>Rapidité de développement</i>	Il est possible de créer une interface fonctionnelle en quelques heures seulement.
<i>Déploiement simplifié</i>	Compatible avec Streamlit Cloud, Heroku, et autres services de déploiement, ce qui facilite la mise en ligne de l'application.

Tableau III-3 : Critères et Justifications du Choix de Streamlit

Bien qu'**Anaconda** soit largement utilisé dans le milieu académique et industriel pour gérer les environnements virtuels et les paquets Python, nous avons choisi de ne pas l'utiliser dans ce projet. Le tableau (III-4) résume les raisons principales justifiant ce choix.

<i>Raisons</i>	<i>Explication</i>
<i>Environnement préconfiguré sur Kaggle</i>	Kaggle fournit nativement un environnement Python complet avec toutes les bibliothèques nécessaires, rendant inutile l'installation manuelle d'Anaconda.
<i>Gain de temps et simplicité</i>	L'utilisation d'Anaconda impliquerait une configuration locale supplémentaire, ce qui aurait allongé le processus de développement sans apport réel.
<i>Limitations matérielles locales</i>	Notre machine dispose d'une carte graphique GTX 1650 (4 Go VRAM), ce qui limite les performances en local. Utiliser Anaconda avec TensorFlow-GPU aurait requis une installation plus complexe et risquée.
<i>Limitations matérielles locales</i>	Le recours à Kaggle nous a permis de contourner les contraintes locales et de bénéficier d'un environnement optimisé dès le départ.

Tableau III-4 : Raisons de non-utilisation d'Anaconda

III.3 Base de données

La base de données utilisée dans ce travail s'intitule TN3K (Thyroid Nodule Region Segmentation Dataset). Elle est largement utilisée dans le domaine de la segmentation des images échographiques des nodules thyroïdiens et est disponible publiquement sur la plateforme Kaggle [21].

III.3.1 Présentation générale du jeu de données TN3K

Le jeu de données TN3K, tableau (III-5), a été conçu spécifiquement pour la tâche de segmentation sémantique des nodules thyroïdiens à partir d'images échographiques. Il contient une collection d'images annotées manuellement par des experts médicaux, ce qui en fait une ressource fiable pour l'entraînement et l'évaluation de modèles de deep learning.

<i>Caractéristiques</i>	<i>Description</i>
<i>Nom du dataset</i>	TN3K – Thyroid Nodule Region Segmentation Dataset
<i>Domaine</i>	Imagerie médicale – Échographie thyroïdienne
<i>Type d'images</i>	Échographies en niveaux de gris
<i>Annotations</i>	Masques binaires segmentés manuellement (nodules)
<i>Nombre total d'images</i>	Environ 3000 paires (image + masque)

Tableau III-5 : Caractéristiques principales du dataset TN3K

III.3.2 Structure du jeu de données

Le dataset est organisé en trois ensembles principaux :

- Entraînement (train)
- Validation (validation)
- Test (test)

Chaque ensemble contient deux sous-dossiers : image, contenant les images échographiques et mask, contenant les masques de segmentation associés. Les images sont au format « .png », en niveaux de gris, tandis que les masques sont des images binaires où chaque pixel est

étiqueté soit comme appartenant au nodule (valeur = 1), soit comme fond (valeur = 0).

III.4 Prétraitement des images

Le prétraitement des images constitue une étape essentielle dans toute chaîne de traitement d'images, notamment dans le cadre de l'apprentissage automatique et du Deep Learning. Il consiste à transformer les données brutes en un format adapté à l'entrée du modèle, tout en améliorant la qualité visuelle et la cohérence des images. Cette phase vise à :

- Uniformiser les dimensions et les formats,
- Réduire le bruit et améliorer le contraste,
- Normaliser les intensités pixel par pixel,
- Faciliter la convergence du modèle lors de l'entraînement.

Dans notre cas, les images échographiques étant souvent sujettes à des problèmes de faible contraste, de bruit et de variations d'intensité dues aux conditions d'acquisition, un ensemble de techniques de prétraitement a été appliqué pour pallier ces défis.

Parmi les différentes techniques de rehaussement d'image, nous avons opté pour l'égalisation d'histogramme. Celle-ci redistribue les intensités des pixels de manière à maximiser le contraste global de l'image, ce qui facilite la détection des contours fins et des structures médicalement pertinentes (comme les nodules thyroïdiens).

L'égalisation d'histogramme consiste à transformer l'image de sorte que sa distribution d'intensité devienne uniforme. Elle utilise la fonction cumulative (eq.III.1) de la distribution initiale des niveaux de gris pour mapper les valeurs vers une plage plus large et mieux distribuée.

$$sk = T(rk) = (L - 1) / N * \sum (j = 0 \text{ à } k) nj$$

Où :

rk : niveau de gris initial.

sk : niveau de gris après transformation.

nj : nombre de pixels ayant l'intensité j.

N : nombre total de pixels.

L : nombre de niveaux de gris (généralement 256).

Pour d'illustrer visuellement l'impact de l'égalisation d'histogramme sur la qualité des images échographiques, la figure (III-1) est une représentation comparative montrant une image originale avant et après traitement.

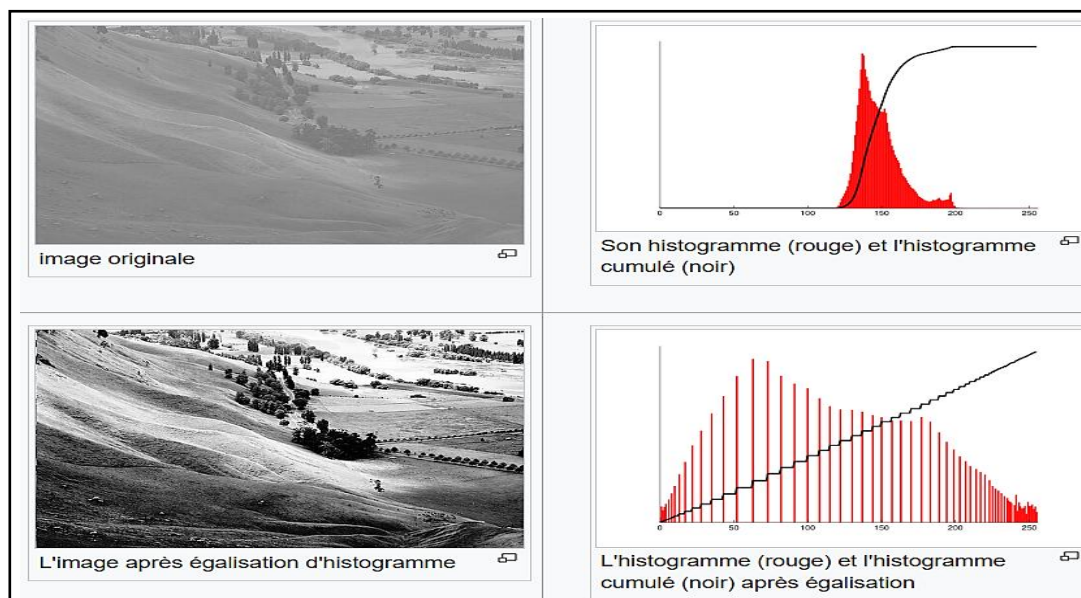


Figure III-1 : Effet de l'égalisation d'histogramme sur une image échographique

Afin de synthétiser les différentes opérations appliquées lors du prétraitement des images échographiques, le tableau (III-6) présente un aperçu global des étapes réalisées ainsi que leurs objectifs respectifs.

<i>Architecture</i>	<i>Avantages</i>	<i>Inconvénients</i>
<i>Redimensionnement</i>	Passage à une taille fixe (256 × 256 pixels)	Uniformiser les dimensions d'entrée pour l'architecture U-Net
<i>Normalisation</i>	Passage des valeurs [0–255] → [0–1]	Stabiliser l'entraînement du réseau et accélérer la convergence
<i>Égalisation d'histogramme</i>	Amélioration du contraste global via la redistribution des intensités.	Faciliter la détection des contours fins et renforcer la visibilité des nodules

Tableau III-6 : Tableau récapitulatif des étapes de prétraitement

III.5 Modèle de Segmentation

Le modèle de segmentation adopté dans cette étude est U-Net, une architecture de réseau de neurones convolutifs conçue spécifiquement pour les applications biomédicales. Elle est particulièrement efficace pour les tâches de segmentation sémantique pixel-par-pixel, où la précision des contours et la localisation des structures anatomiques sont cruciales [22].

III.5.1 Architecture générale d’U-Net

L’architecture d’U-Net se compose de deux chemins principaux :

- **Un chemin d’encodage (contractant)** : composé de plusieurs blocs de couches Conv2D, BatchNorm, ReLU, MaxPooling, permettant d’extraire des caractéristiques hiérarchiques tout en réduisant la dimension spatiale.
- **Un chemin de décodage (expansif)** : constitué de blocs comprenant un upsampling, concaténation avec les sorties correspondantes du chemin d’encodage, Conv2D, BatchNorm, ReLU, visant à restaurer les dimensions spatiales tout en préservant l’information fine via les skip connections.

Enfin, une couche Conv2D avec une fonction d’activation sigmoïde produit une carte de segmentation binaire où chaque pixel est classé comme appartenant au nodule ou au fond.

III.5.2 Caractéristiques clés

Dans le tableau (III-7), nous détaillons les caractéristiques du modèle U-Net comme suite :

<i>Caractéristiques</i>	<i>Description</i>
Structure symétrique	L’architecture contractante/expansive assure une reconstruction précise des contours.
Skip connections	Permettent de combiner les informations spatiales du début du réseau avec les représentations sémantiques profondes, améliorant ainsi la qualité de la segmentation.
Robustesse aux données limitées	U-Net peut être entraîné efficacement même avec peu de données annotées, ce qui est crucial dans les applications médicales.

<i>Précision sur petites structures</i>	Idéal pour la détection de petits objets tels que les nodules thyroïdiens.
--	--

Tableau III-7 : Caractéristiques principales d’U-Net

III.5.3 Justification du choix

Les images échographiques présentent souvent un faible contraste, des artefacts et une texture hétérogène. U-Net s’adapte bien à ces défis grâce à :

- Une meilleure localisation spatiale due aux skip connections,
- Une capacité à généraliser malgré le bruit,
- Une structure modulaire facile à adapter à différents types de données.

De plus, dans notre cas spécifique, l’utilisation de masques binaires rend U-Net particulièrement pertinent, car il est capable de produire des segmentations précises avec une sortie probabiliste.

III.5.4 Comparaison avec d'autres architectures

Comparé à d’autres architectures convolutionnelles telles que VGG, Inception ou DeepLabv3+, U-Net offre plusieurs avantages cités dans le tableau (III-8) suivant:

<i>Architecture</i>	<i>Avantages</i>	<i>Inconvénients</i>	<i>Pertinence pour TN3K</i>
<i>U-Net</i>	Précision, simplicité, rapidité	Moins adapté aux textures complexes	Adopté, performances optimales
<i>DeepLabv3+</i>	Bonne gestion multi-échelle	Plus complexe, nécessite GPU puissant	Non utilisé
<i>Attention U-Net</i>	Mécanismes d’attention pour meilleure localisation	Complexité accrue	Option intéressante mais non testée
<i>ResUNet</i>	Blocs résiduels pour convergence rapide	Besoin de plus de données et de ressources	Option potentielle pour futurs travaux

MultiResUNet	Filtrage multi-résolution	Coût computationnel élevé	Peut améliorer les résultats mais non testé
---------------------	---------------------------	---------------------------	---

Tableau III-8 : Comparaison des architectures basées sur U-Net pour la segmentation des nodules thyroïdiens

Cependant, malgré les améliorations proposées par ces variantes, **les résultats** expérimentaux préliminaires sur notre base de données TN3K ont montré que l’U-Net classique offrait un compromis optimal entre précision, rapidité d’entraînement et complexité computationnelle. En effet, il a permis d’atteindre un Dice coefficient moyen de ~ 0.82 sur le jeu de test, sans nécessiter de matériel GPU haut de gamme.

III.6 Fonction d’activation

Les fonctions d’activation jouent un rôle fondamental dans les réseaux de neurones profonds. Elles introduisent de la non-linéarité dans le modèle, permettant ainsi au réseau de capturer des relations complexes entre les données d’entrée et les sorties attendues. Sans ces fonctions, un réseau de neurones se réduirait à une simple combinaison linéaire, incapable de modéliser des tâches avancées comme la segmentation sémantique.

Dans notre implémentation du modèle U-Net, deux fonctions d’activation principales ont été utilisées :

- **ReLU (Rectified Linear Unit) :**

Appliquée dans toutes les couches cachées du réseau, et définie par la fonction représentée par l’équation (III.2) :

$$\text{ReLU}(x) = \max(0, x)$$

Elle présente plusieurs avantages :

- Évite le problème de saturation des gradients par rapport aux fonctions comme sigmoïde.
- Accélère la convergence lors de l’entraînement,
- Favorise la sparsité du réseau en annulant les valeurs négatives.

- **Sigmoïde :**

Appliquée uniquement à la couche de sortie pour produire une carte de segmentation binaire, et définie par la fonction représentée par l'équation (III.3) :

$$\sigma(x) = 1 / (1 + e^{(-x)})$$

Elle présente plusieurs avantages :

- Retourne des valeurs comprises entre 0 et 1, interprétables comme des probabilités.
- Adaptée à la segmentation binaire où chaque pixel doit être classé comme appartenant au nodule (1) ou au fond (0).

Bien que cette fonction puisse souffrir de saturation des gradients, son utilisation reste pertinente à la couche de sortie pour une tâche de segmentation avec masque binaire.

III.7 Fonction de perte

La fonction de perte (ou loss function) mesure l'écart entre les prédictions du modèle et les valeurs réelles (vérité terrain). Elle guide l'optimisateur pendant l'entraînement pour ajuster les poids du réseau afin de minimiser cet écart. Deux fonctions de perte principales ont été combinées dans notre travail :

- **Dice Loss :**

Basée sur le coefficient de Dice, elle mesure la similarité entre la prédiction et la vérité terrain, et définie par la fonction représentée par l'équation (III.4) :

$$\text{Dice Loss} = 1 - \frac{2 \cdot |A \cap B| + \epsilon}{|A| + |B| + \epsilon}$$

Où A est la prédiction et B la vérité terrain, ϵ un terme de lissage pour éviter la division par zéro. Il est très efficace pour les jeux de données déséquilibrés, comme c'est le cas ici (beaucoup plus de pixels de fond que de pixels de nodule), et idéale pour la segmentation médicale où les structures d'intérêt sont souvent petites et rares [23].

- **Binary Cross-Entropy (BCE) :**

Mesure la divergence entre deux distributions binaires (prédiction vs réalité), et définie par la

$$\text{BCE} = -\frac{1}{N} \sum_{i=1}^N y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$$

fonction représentée par l'équation (III.5) :

Il stabilise l'entraînement en renforçant la précision globale du modèle, souvent combinée à la Dice Loss pour compenser ses faiblesses potentielles, notamment dans la gestion des erreurs localisées.

Dans notre implémentation, nous avons adopté une combinaison linéaire de Dice Loss et Binary Cross-Entropy, ce qui permet d'atteindre un bon compromis entre stabilité d'entraînement, gestion du déséquilibre des classes et précision spatiale. La formule combinée est représentée par l'équation (III.6) :

$$\text{Total Loss} = \alpha \cdot \text{Dice Loss} + (1 - \alpha) \cdot \text{BCE}$$

où $\alpha \in [0,1]$ est un paramètre de pondération ajustable.

III.8 Modèle de Classification

En complément de la segmentation, la classification automatique des nodules thyroïdiens permet de prédire la nature bénigne ou maligne des nodules détectés dans les images échographiques. Cette tâche de classification s'inscrit dans une logique de soutien au diagnostic médical, en réduisant le recours aux biopsies invasives. Dans ce projet, nous avons mis en place un pipeline de classification basé sur des réseaux de neurones convolutifs (CNN), entraînés sur un dataset limité d'images annotées [24].

III.8.1 Architecture du modèle de classification

Le modèle utilisé repose sur une architecture CNN classique, constituée de plusieurs couches convolutives, de normalisation, de pooling et de couches entièrement connectées

(Dense), suivies d'une sortie binaire avec une activation sigmoïde. Le tableau (III-9) suivant résume la structure du modèle.

Couche	Paramètres principaux
Conv2D	32 filtres, noyau 3x3, activation ReLU
MaxPooling2D	Pool size 2x2
Dropout	0.25
Conv2D	64 filtres, noyau 3x3, activation ReLU
MaxPooling2D	Pool size 2x2
Flatten	-
Dense	128 unités, activation ReLU
Dropout	0.5
Dense (Sortie)	1 unité, activation sigmoïde

Tableau III-9 : structure du modèle CNN

La couche convolutionnelle 2D (Conv2D) : est l'un des composants les plus fondamentaux et efficaces des réseaux de neurones convolutifs (CNN) pour le traitement des données d'image. Elle est particulièrement bien adaptée à l'analyse des images échographiques médicales, comme celles des nodules thyroïdiens, grâce à sa capacité à extraire automatiquement des caractéristiques discriminantes locales [25].

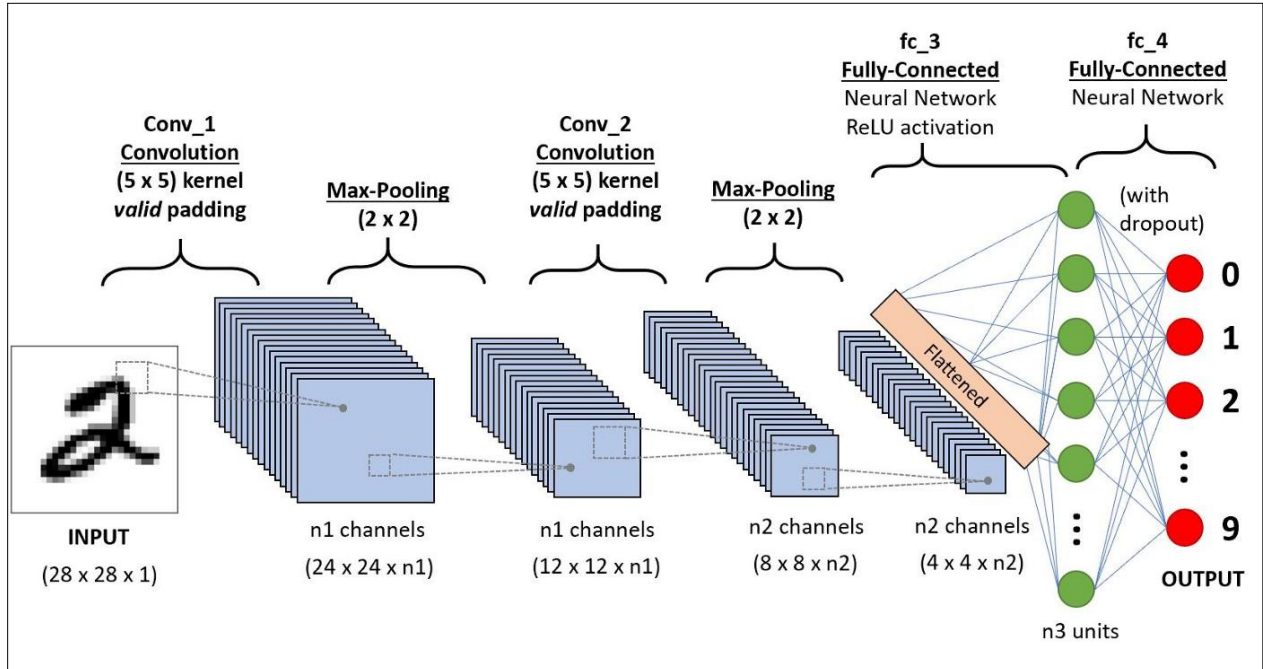


Figure III-2 : Architecture d'un CNN

La couche Conv2D, ainsi illustrée dans la figure (III-2), applique une série de filtres (ou noyaux) de petite taille (souvent 3×3) qui glissent sur l'image (opération de convolution) pour détecter des motifs tels que des bords, des textures, des contours, des formes [26].

La figure (III-3) montre un exemple de convolution.

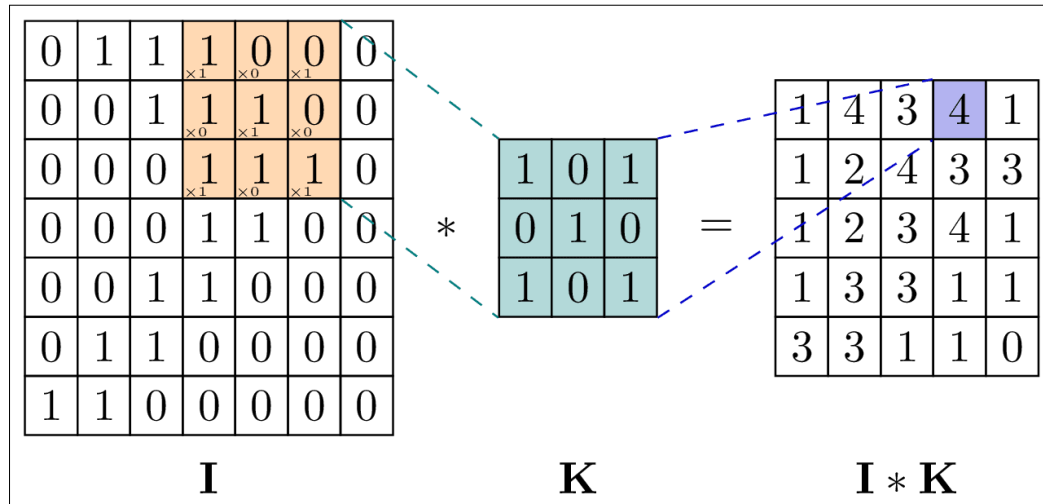


Figure III-3 : couche de convolution

Avec

I : est une matrice qui représente l'image d'entrée ou la carte de caractéristiques issue d'une couche précédente (input).

K : est un filtre (ou noyau de convolution) appliqué à l'image.

I*k : Résultat de la convolution (output).

En imagerie échographique thyroïdienne, les couches Conv2D sont particulièrement adaptées pour détecter :

- les bords irréguliers des nodules,
- les textures internes des masses,
- les calcifications ou zones anéchogènes,
- les caractéristiques échogènes hétérogènes associées à la malignité.

Le Pooling : est une opération fondamentale dans les réseaux de neurones convolutifs (CNN), utilisée pour réduire les dimensions spatiales (hauteur et largeur) des cartes de caractéristiques tout en conservant les informations les plus significatives. Il agit comme une forme de résumé local des données extraites par les couches convolutionnelles [27].

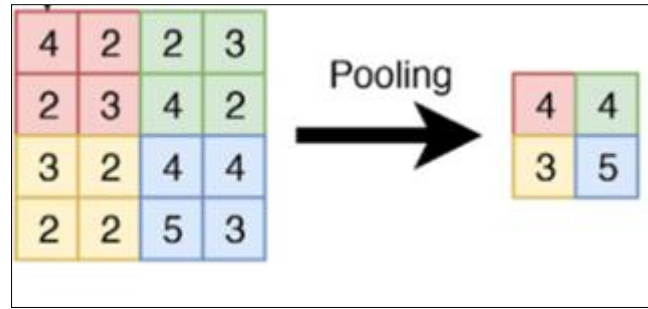


Figure (III-4) : Représentation schématique de l'opération de pooling dans un CNN.

Le pooling agit sur des sous-régions locales (généralement 2×2 ou 3×3). Il résume cette région à une seule valeur selon une règle spécifique, afin de réduire la taille des cartes, donc moins de calculs, prévenir un surapprentissage en supprimant des détails trop fins, et aussi garantir une robustesse aux translations mineures dans l'image [27].

Dans les images échographiques, par exemple pour les nodules thyroïdiens, le pooling aide à :

- Conserver les textures dominantes,
- Éliminer le bruit échographique,
- Renforcer la structure des bords ou calcifications importantes [25].

III.9 Conclusion

Ce chapitre a permis de définir le cadre technique et méthodologique retenu pour la mise en œuvre du modèle de segmentation U-Net et de classification CNN. Le choix du matériel, de l'environnement de simulation (Kaggle), ainsi que l'utilisation du dataset TN3K ont été justifiés en fonction des contraintes pratiques et des objectifs visés. L'architecture U-Net, associée à un prétraitement adapté des images et à une combinaison efficace de fonctions de perte, s'est révélée bien adaptée à la tâche de segmentation des nodules thyroïdiens. L'architecture CNN classique, combinant des couches convolutionnelles, de pooling, de régularisation (Dropout) et des couches entièrement connectées, est la méthode de choix pour la classification automatique des nodules thyroïdiens. Ces fondations solides permettent d'aborder dans de bonnes conditions l'évaluation des résultats présentée dans le chapitre suivant.

CHAPITRE IV.

Résultats expérimentaux et évaluation du modèle

Chapitre IV . Résultats expérimentaux et évaluation du modèle

IV.1 Introduction

Après avoir détaillé dans le chapitre précédent l’environnement de développement, les caractéristiques du jeu de données TN3K, ainsi que l’architecture U-Net choisie pour la segmentation des nodules thyroïdiens, nous abordons à présent la phase d’évaluation expérimentale.

Ce chapitre présente les résultats obtenus suite à l’entraînement du modèle, en s’appuyant sur des métriques quantitatives (telles que le Dice coefficient et l’IoU) ainsi que des visualisations qualitatives des masques prédits. Il met également en évidence l’évolution de la performance du modèle au fil des époques, tout en discutant des techniques d’amélioration appliquées pour optimiser la convergence et la qualité des prédictions. Ces résultats permettent non seulement de valider la pertinence de l’approche proposée, mais aussi de la comparer à d’autres architectures de segmentation afin d’évaluer sa robustesse et sa capacité de généralisation.

IV.2 Choix de jeu de données

Le jeu de données TN3k se distingue parmi les bases de données disponibles sur Kaggle pour la segmentation des nodules thyroïdiens, et le tableau ci-dessus représente une analyse comparative mettant en évidence ses avantages par rapport à d’autres jeux de données similaires.

<i>Critère</i>	<i>TN3K</i>	<i>DDTI</i>	<i>TNUI-2021</i>
<i>Nombre d’images</i>	3 493	1 000	1 000
<i>Nombre de patients</i>	2 421	Non spécifié	Non spécifié
<i>Période de collecte</i>	Janvier 2016 – Août 2020	Non spécifiée	Non spécifiée
<i>Type d’images</i>	Échographies 2D en niveaux de gris	Échographies 2D en niveaux de gris	Échographies 2D en niveaux de gris

Annotations	Masques manuels de haute qualité	Masques manuels	Masques manuels
Diversité des nodules	Large variété de tailles et de formes	Moins de diversité	Moins de diversité
Accessibilité	Kaggle	Kaggle	GitHub
Utilisation dans la recherche	Largement utilisé dans des études récentes pour l'évaluation de modèles avancés de segmentation	Utilisé dans des études antérieures	Moins utilisé dans la littérature récente

Tableau IV-1 : Comparaison des jeux de données pour la segmentation des nodules

IV.2.1 Avantages du jeu de données TN3K :

- **Volume de données conséquent :** avec 3493 images provenant de 2421, TN3K offre une base solide pour l'entraînement de modèles de Deep Learning réduisant ainsi le risque de surapprentissage.
- **Annotation de haute qualité :** les masques de segmentations ont été réalisés manuellement assurant une précision élevée des annotations, ce qui est crucial pour l'entraînement et l'évaluation des modèles.
- **Diversité des cas cliniques :** La base couvre une large gamme de tailles, formes et textures de nodules, reflétant la variabilité clinique réelle et améliorant la robustesse des modèles développés.
- **Utilisation dans des études récentes :** TN3K a été employé dans des recherches récentes pour évaluer des architectures avancées telles qu'Improver U-Net démontrant sa pertinence et sa fiabilité dans le domaine.
- **Accessibilité et documentation :** Disponible sur kaggle avec une documentation claire.

Alors, le jeu de données TN3k se distingue par sa taille, la qualité de ses annotations et sa diversité, en faisant une ressource précieuse pour le développement et l'évaluation de modèles de segmentation des nodules thyroïdiens. Par rapport à d'autres bases de données disponibles sur kaggle, Tn3K offre des avantages significatifs qui peuvent contribuer à des avancées dans le diagnostic assisté par ordinateur des maladies thyroïdiennes.

IV.3 Analyse comparative des modèles et justification du choix final pour la segmentation

IV.3.1 Expérimentation 1 « 1^{ère} évaluation avec le modèle DeepLabv3+ »

Comme première expérimentation de segmentation, nous avons mis en œuvre le modèle DeepLabv3+ avec ResNet50 (hybride) comme backbone, en l'ajustant selon les paramètres présentés dans le tableau ci-dessus.

<i>Paramètre</i>	<i>Valeurs</i>
Architecture	DeepLabv3+ avec ResNet50 backbone
Nombre d'époques (epochs)	30
Batch size	8
Taux d'apprentissage	0.001
Fonction de perte	BCEWithLogitsLoss
Optimiseur	Adam
Scheduler	ReduceLROnPlateau
Couche finale modifiée	Conv2d(256, 1, kernel_size=1)
Fonction d'activation finale	Sigmoid (via torch.sigmoid en posttraitement)
Temps d'entraînement	3h42min

Tableau IV-2 : Paramètres d'entraînement du modèle DeepLabv3+

L'entraînement du modèle DeepLabv3+ avec un backbone ResNet50 sur le dataset échographique TN3K a montré des courbes d'apprentissage (loss, Dice, accuracy), représentées dans la figure suivante, très stables et ascendantes, suggérant a priori une bonne

capacité d'apprentissage. En particulier, les courbes de Dice coefficient et de précision pixel en validation affichent une progression régulière, sans surapprentissage manifeste sur les 30 époques.

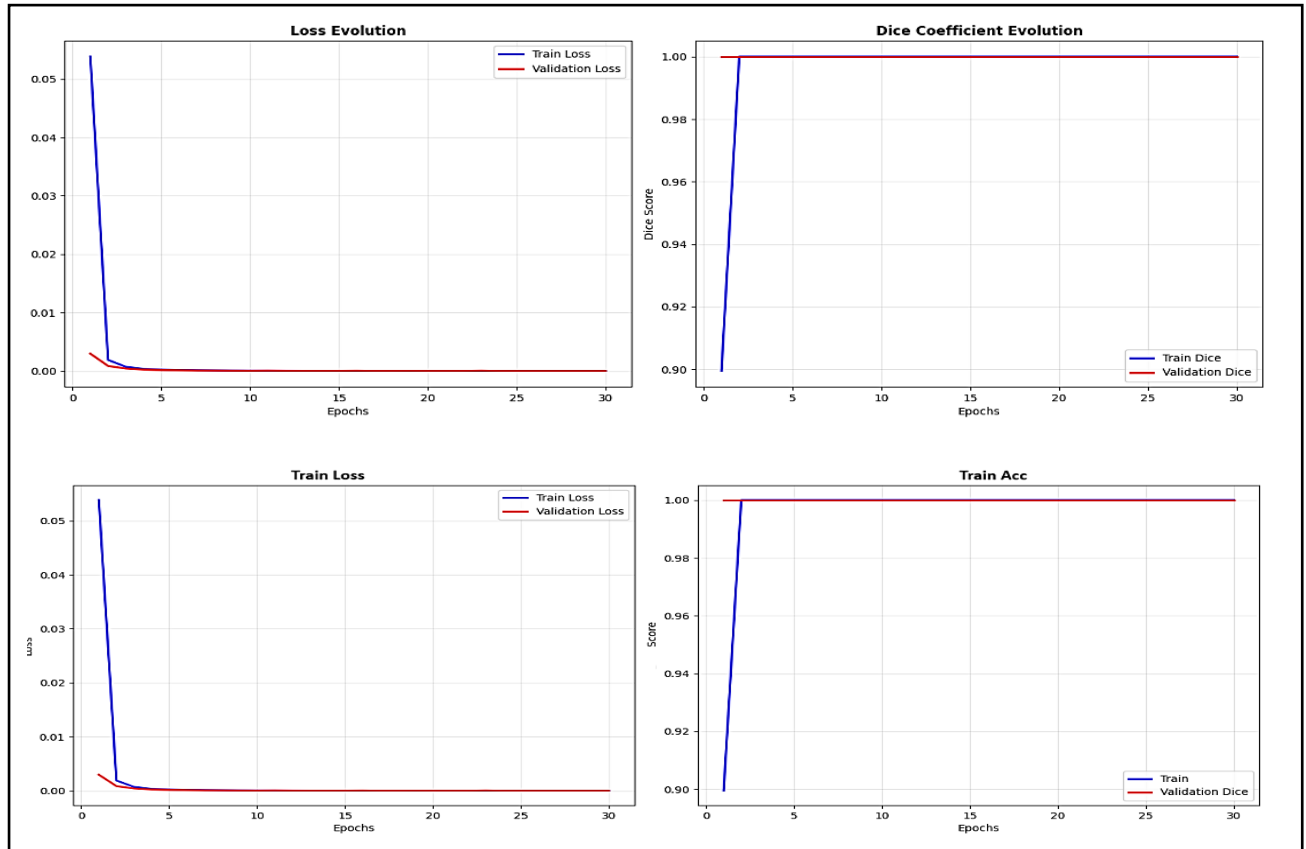


Figure IV-1 : Courbes d'apprentissage DeepLabv3+.

Cependant, une analyse visuelle, figure IV-2, des masques prédits révèle un écart significatif entre les performances mesurées et les résultats réels de segmentation. Malgré des scores élevés sur les métriques classiques, le modèle ne parvient pas à correctement segmenter les nodules :

- Les masques prédits sont absents ou mal alignés avec les structures anatomiques réelles
- Certaines images présentent des prédictions quasi nulles
- Le modèle ignore totalement le masque réel, même après convergence apparente du loss.

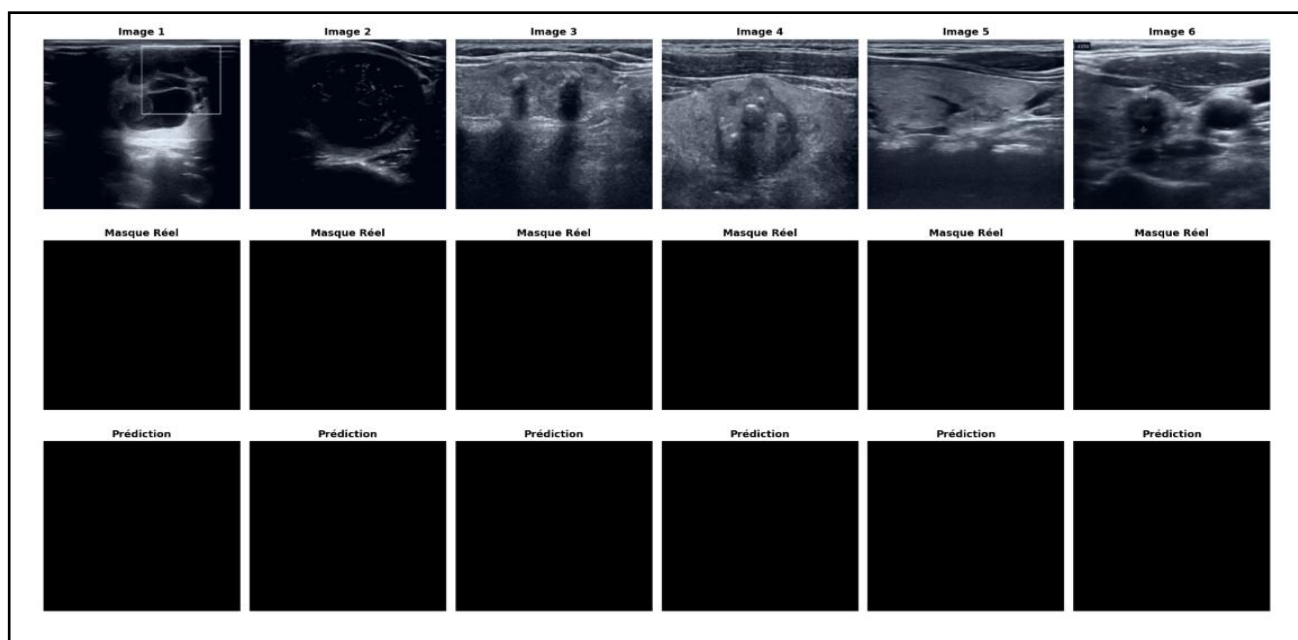


Figure IV-2 : résultats de segmentation du modèle DeepLabv3+.

Ce décalage entre les métriques quantitatives et la qualité visuelle des résultats indique une forme de sur apprentissage sur des motifs non pertinents, ou une mauvaise généralisation aux structures fines et floues caractéristiques des nodules thyroïdiens. Plusieurs facteurs explicatifs peuvent être envisagés :

- Architecture trop complexe pour le volume de données disponible : DeepLabv3+ est conçu pour des scènes larges avec objets bien délimités, pas pour des objets petits et faiblement contrastés.
- Perte de détails fins due à l'absence de skip connections
- Pas de mécanismes explicites pour forcer la reconnaissance de contours ou de textures faibles, essentiels en échographie.

Bien que les indicateurs numériques soient rassurants, le comportement réel du modèle sur les images test remet en question sa pertinence pour ce type de tâche médicale. Cette expérimentation met donc en évidence la limite d'interprétation des courbes de performance sans analyse qualitative, et justifie l'orientation vers une architecture plus adaptée aux contraintes médicales, mieux armée pour préserver les détails morphologiques fins grâce à ses skip connections.

IV.3.2 Expérimentation 2 « 2^{ème} évaluation avec le modèle YOLOv5+ XGBoost »

Dans notre deuxième expérimentation de segmentation, nous avons mis en œuvre le modèle **YOLOv5 avec XGBoost**, en l'ajustant selon les paramètres présentés dans le tableau ci-dessus.

<i>Paramètre</i>	<i>Valeurs</i>
Architecture	YOLOv5s
Nombre d'époques (epochs)	50
Batch size	16
Taux d'apprentissage	Adaptatif (géré automatiquement)
Fonction de perte	CIoU + Binary Cross-Entropy
Optimiseur	Adam
Scheduler	ReduceLROnPlateau
Couche finale modifiée	Aucune
Fonction d'activation finale	Sigmoid
Temps d'entraînement	Non terminé (ressources insuffisantes)

Tableau IV-3 : Paramètres d'entraînement du modèle YOLOv5+XGboost.

Dans cette expérimentation, l'objectif était d'exploiter le modèle **YOLOv5+GXboots** pour la détection automatique des nodules thyroïdiens à partir des images échographiques. Le pipeline prévoyait un entraînement sur 50 époques à une résolution de 640×640, avec un batch size de 16. L'entraînement a été lancé dans l'environnement Kaggle, équipé d'un GPU P100.

Bien que l'entraînement ait démarré correctement, les valeurs de perte restent élevées et les métriques de performance (mAP) indiquent une précision encore très faible. La mémoire GPU

atteint rapidement ses limites (~16 Go), ce qui ralentit considérablement la simulation et

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size		
1/50 0:00, 3.41it/s]	4.13G	1.623	2.127	1.755	31	640: 100% ██████████ 144/144 [00:42<0		
[00:05<00:00, 3.40it/s]	Class	Images	Instances	Box(P	R	mAP50	mAP50-95): 100% ██████████ 18/18	
	all	576	576	0.281	0.368	0.234	0.0885	
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size		
2/50 0:00, 3.65it/s]	4.17G	1.685	1.801	1.806	36	640: 100% ██████████ 144/144 [00:39<0		
[00:03<00:00, 4.53it/s]	Class	Images	Instances	Box(P	R	mAP50	mAP50-95): 100% ██████████ 18/18	
	all	576	576	0.331	0.325	0.241	0.0835	
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size		
3/50 0:00, 3.69it/s]	4.2G	1.674	1.766	1.798	30	640: 100% ██████████ 144/144 [00:38<0		
[00:03<00:00, 4.64it/s]	Class	Images	Instances	Box(P	R	mAP50	mAP50-95): 100% ██████████ 18/18	
	all	576	576	0.416	0.364	0.323	0.131	

⚠ Warning about the use of GPU

risque de provoquer une interruption anticipée.

Figure IV-3 : Résultat de simulation de modelé YOLOv5+GXboots sur kaggle

Malgré une préparation complète du pipeline, l'entraînement du modèle YOLOv5 n'a pas pu aboutir en raison de plusieurs contraintes, comme la Taille et complexité du modèle YOLOv5, les Exigences matérielles élevées ainsi que son Temps d'entraînement excessif.

Cette tentative d'utilisation de YOLOv5 pour la détection des nodules échographiques a montré ses limites pratiques dans un environnement de ressources restreintes. Elle met en évidence l'importance d'adapter le choix du modèle aux contraintes matérielles disponibles, surtout en traitement médical où la légèreté et l'efficacité sont essentielles.

Pour les prochaines expérimentations, nous nous orienterons vers des architectures plus simples, modulaires et compatibles avec des environnements limités, comme Attention U-Net ou U-Net classique.

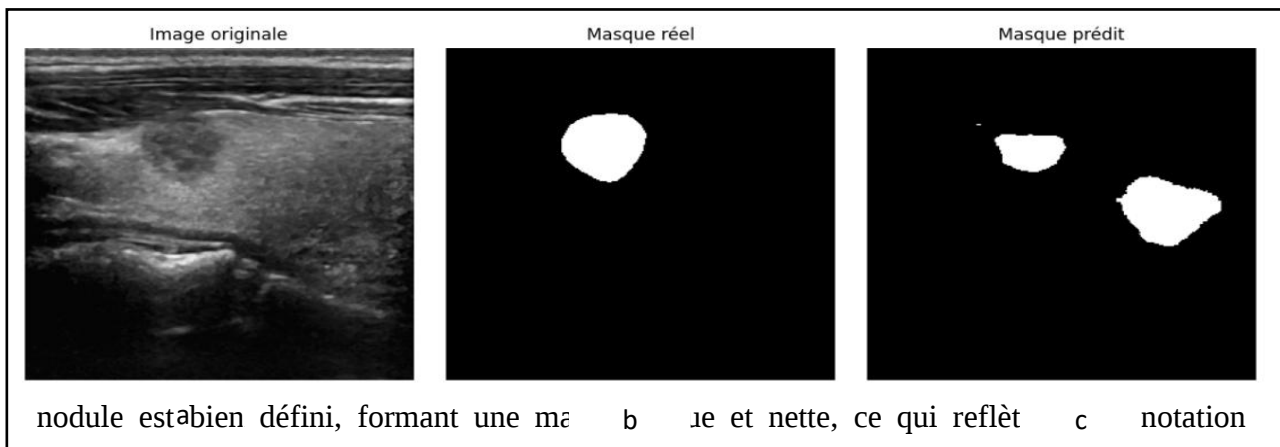
IV.3.3 Expérimentation 3 « 3^{ème} évaluation avec le modèle Attention U-Net »

Dans cette troisième expérimentation de segmentation, nous avons implémenté le modèle Attention U-Net, en l'entraînant selon les paramètres récapitulés dans le tableau ci-dessus.

<i>Paramètre</i>	<i>Valeurs</i>
Architecture	AttentionU-Net
Nombre d'époques (epochs)	40
Batch size	8
Taux d'apprentissage	0.001
Fonction de perte	Binary Cross-Entropy + Dice loss
Optimiseur	Adam
Scheduler	ReduceLROnPlateau
Couche finale modifiée	Conv2D (1 canal, sigmoïde)
Fonction d'activation finale	Sigmoid
Temps d'entraînement	52 min

Tableau IV-4 : Paramètres d'entraînement du modèle Attention U-Net.

Sur la figure (IV-a), nous observons un nodule bien visible au centre de l'image échographique. En plus, le masque réel est bien illustre sur la figure (IV-b), ou Le contour du



manuelle précise. Pour ce qui est du Masque prédit par le modèle, figure (IV-c), nous remarquons que le modèle a réussi à détecter deux masses séparées, au lieu d'une seule. Cela indique un problème de cohérence structurelle dans la prédiction, et le modèle tend à sursegmenter, en détectant des régions bruitées comme étant des nodules. En fin que Le positionnement général est relativement correct, mais la forme et la connectivité ne correspondent pas au masque réel. Ce type de sortie révèle une bonne sensibilité (le modèle capte des structures) mais une précision limitée (faible exactitude des formes). Cela explique un Dice coefficient de validation modéré (~ 0.79) malgré une haute accuracy : le modèle prédit trop, mais pas de manière ciblée.

Figure IV-4 : résultats de segmentation du modèle Attention U-Net.

Les courbes, figure suivante révèlent une forte tendance au surapprentissage à partir de la 20 époque. Ainsi, le modèle ne parvient pas à généraliser correctement, malgré un apprentissage réussi sur les données d'entraînement. Cela justifie la nécessité de tester d'autres architectures, plus simples ou mieux adaptées à la nature des images échographiques.

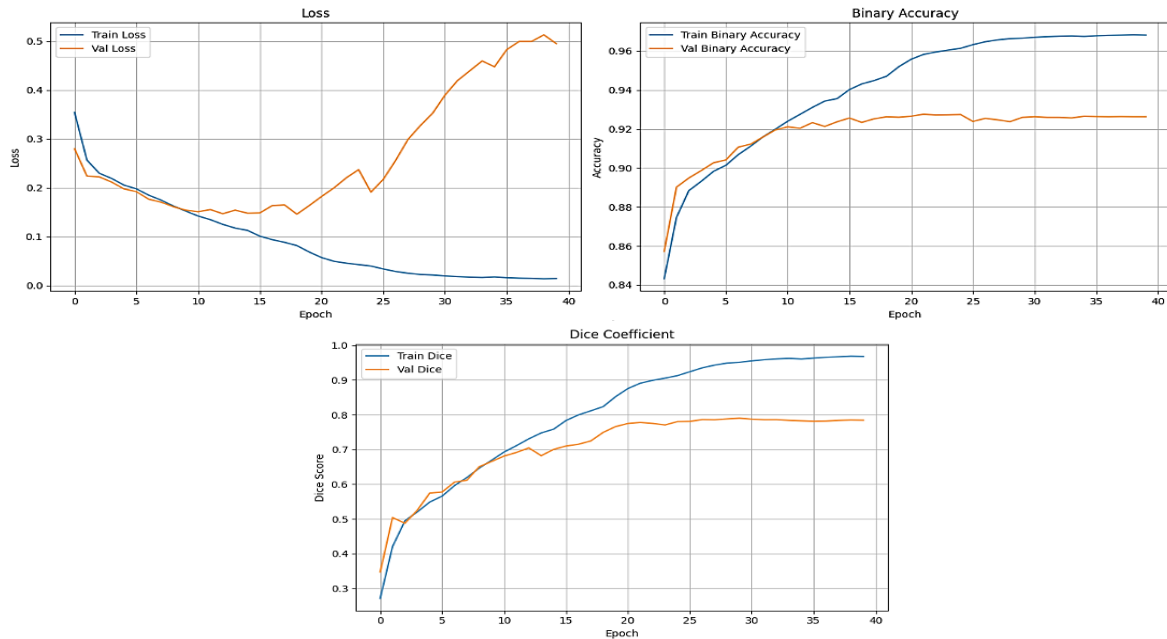


Figure IV-5 : Courbes d'apprentissage Attention U-Net.

Test Binary Accuracy: 0.9458
Test Dice Coefficient: 0.7938

Figure IV-6 : Précision du modèle Attention U-Net.

IV.3.4 Expérimentation 4 « Évaluation du modèle U-Net avec techniques d'amélioration »

Dans cette dernière expérimentation, nous avons exploité pleinement le potentiel du modèle U-Net classique en l'entraînant sur 50 époques, avec une supervision par ReduceLROnPlateau et EarlyStopping, dans l'objectif d'améliorer la précision de la segmentation des nodules thyroïdiens.

Les courbes, figures suivantes, et les métriques témoignent d'un apprentissage stable et efficace, et l'absence d'overfitting est confirmée par la stabilité des courbes de validation. En plus, Dice élevé (> 0.90 en validation) reflète une bonne précision sur la forme des nodules segmentés.

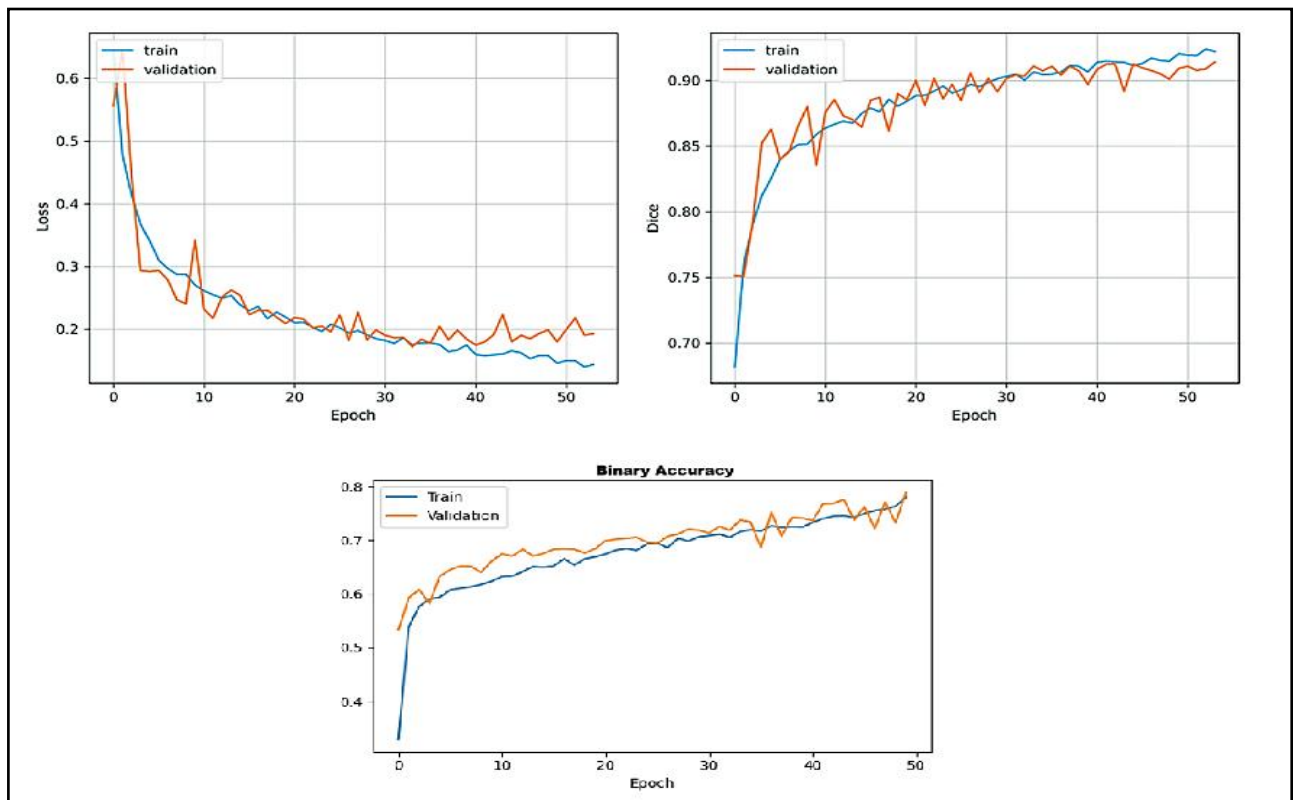


Figure IV-7 : Courbes d'apprentissage du modèle U-Net.

Le modèle U-Net entraîné pendant 50 époques, avec régularisation et ajustement dynamique du taux d'apprentissage, montre des résultats supérieurs à toutes les autres architectures testées (DeepLabv3+, Attention U-Net, YOLOv5), combinant ainsi une excellente convergence des courbes d'apprentissage, et une qualité visuelle supérieure des masques. Il est donc retenu comme modèle final pour la segmentation des nodules thyroïdiens.

IV.4 Étapes de réalisation du modèle de segmentation final

La méthodologie adoptée pour la mise en œuvre du modèle final de segmentation des nodules thyroïdiens repose sur une approche structurée en plusieurs étapes clés. Ce processus vise à exploiter pleinement les capacités des réseaux de neurones convolutifs profonds, en particulier l'architecture U-Net, afin d'obtenir une segmentation précise et fiable des régions d'intérêt dans les images échographiques. La figure IV-7 présente de manière organisée les différentes phases de développement, d'entraînement et d'évaluation du modèle proposé.

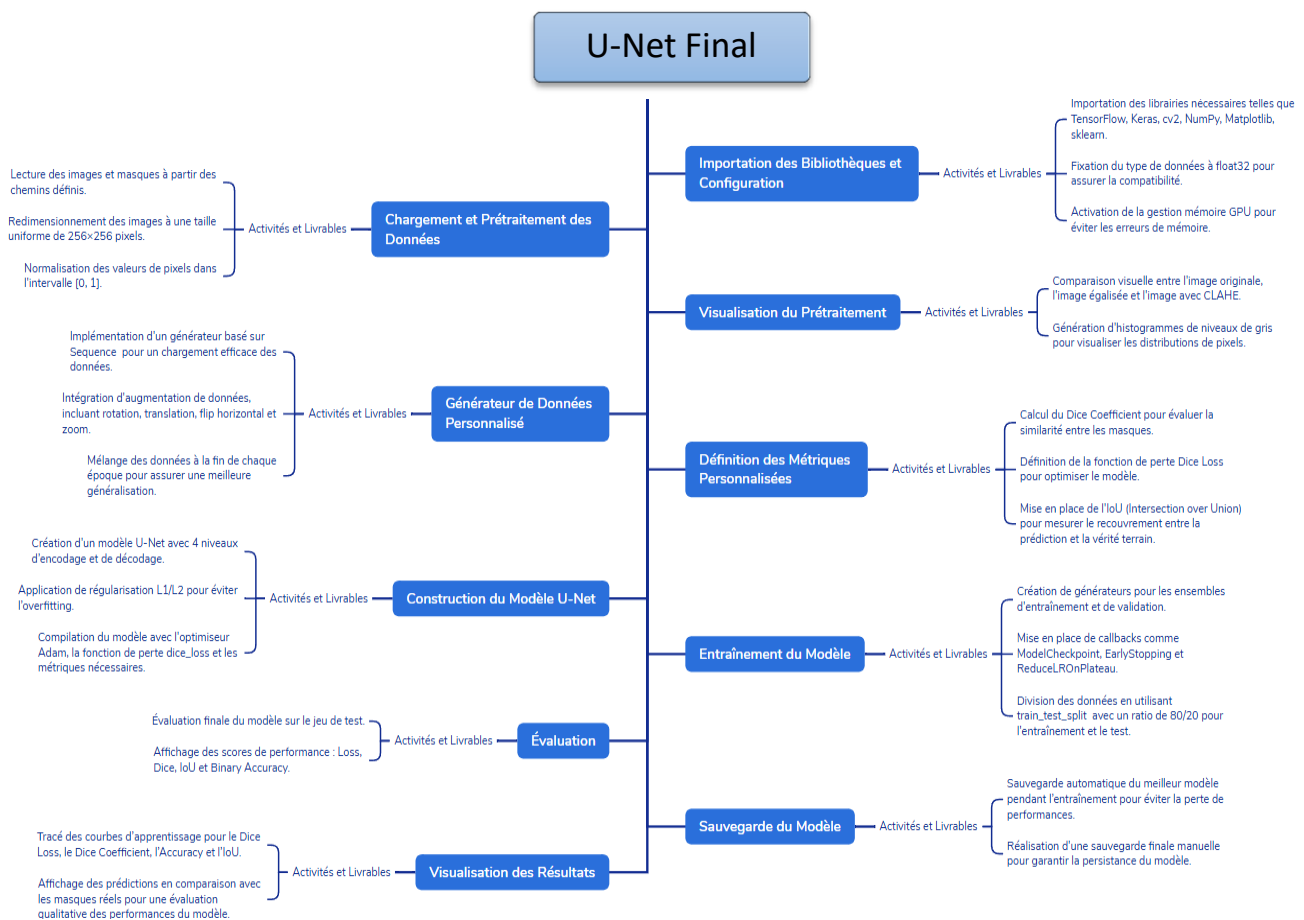


Figure IV-8 : Organigramme méthodologique du modèle U-Net

Les principales étapes de ce processus méthodologique sont les suivantes :

- **Importation des bibliothèques**

Cette étape permet de charger l'ensemble des modules indispensables au traitement des données, à la conception du modèle de Deep Learning, ainsi qu'à son entraînement, son évaluation et à la visualisation des performances obtenues.

- **Chargement et prétraitement des données**

Elle consiste à lire les images et leurs masques correspondants, à les redimensionner à une taille uniforme, à normaliser les pixels, et à améliorer leur qualité visuelle via des techniques comme l'égalisation d'histogramme ou le CLAHE. Ce traitement prépare les données à être exploitées efficacement par le modèle.

- **Visualisation du prétraitement**

Cette étape permet de comparer visuellement les images originales et celles prétraitées. Elle inclut également la génération d'histogrammes pour analyser la distribution des niveaux de gris, facilitant ainsi l'évaluation qualitative des effets du prétraitement.

- **Générateur de données personnalisé**

Un générateur de données dynamique est mis en place pour alimenter le modèle U-Net. Il assure le chargement efficace des données par batchs et applique des techniques d'augmentation (rotation, translation, zoom, flip) pour enrichir artificiellement le jeu de données et améliorer la robustesse du modèle.

- **Définition des métriques personnalisées**

Cette phase vise à définir les fonctions d'évaluation adaptées à la tâche de segmentation. Des métriques comme le Dice Coefficient et l'IoU (Intersection over Union) sont utilisées pour quantifier la qualité de la prédiction, tandis que la Dice Loss sert de fonction de coût pour optimiser l'apprentissage.

- **Construction du modèle U-Net**

Le modèle est conçu selon l'architecture U-Net, bien adaptée aux tâches de segmentation. Il intègre plusieurs niveaux d'encodage et de décodage, avec une régularisation L1/L2 pour prévenir le sur-apprentissage. L'optimiseur Adam est utilisé, accompagné des métriques définies précédemment.

- **Entraînement du modèle**

Le modèle est entraîné sur les données d'apprentissage en utilisant des générateurs augmentés. Des callbacks comme EarlyStopping, ReduceLROnPlateau et ModelCheckpoint

sont intégrés pour optimiser l'entraînement, prévenir la suradaptation et sauvegarder le meilleur modèle.

- **Évaluation du modèle**

Cette étape consiste à évaluer le modèle final sur un jeu de test indépendant. Les performances sont mesurées à l'aide des métriques définies (Loss, Dice, IoU, Accuracy) afin de vérifier la capacité du modèle à bien généraliser sur des données non vues.

- **Sauvegarde du modèle**

Le meilleur modèle obtenu durant l'entraînement est automatiquement sauvegardé au format .h5. Une sauvegarde finale manuelle est également effectuée pour garantir la conservation du modèle entraîné, en vue d'une réutilisation ou d'un déploiement ultérieur.

- **Visualisation des résultats**

Les courbes d'apprentissage sont générées pour illustrer l'évolution des performances (perte, Dice, IoU, précision) au fil des époques. Des prédictions sont comparées visuellement aux masques réels pour valider qualitativement l'efficacité du modèle.

IV.5 Paramètres du modèle d'apprentissage profond

Le développement et l'entraînement du modèle U-Net reposent sur un ensemble de bibliothèques Python spécialisées dans la manipulation d'images, la conception de modèles de Deep Learning, la gestion des données, ainsi que la visualisation des résultats, cités précédemment dans ce manuscrit.

Tableau IV-5 suivant présente les caractéristiques essentielles du jeu de données utilisé, ainsi que les choix effectués pour le prétraitement et la configuration des données.

<i>Paramètre</i>	<i>Valeur / Description</i>
Origine du dataset	TN3K Thyroid Nodule Region Segmentation Dataset (Kaggle)

Type d'images	Échographiques en niveaux de gris
Format des fichiers	.png / .jpg (images et masques)
Nombre d'images (train + val)	3493 paires image/masque
Taille d'image cible	256 × 256 pixels
Canal utilisé	1 (niveaux de gris)
Normalisation	Oui — valeurs des pixels ramenées dans l'intervalle [0, 1]
Égalisation de contraste	Oui — CLAHE (Contrast Limited Adaptive Histogram Equalization)
Prétraitement des masques	Binarisation des masques à l'aide d'un seuillage (seuil = 0.5)
Augmentation de données	Oui — rotation, translation, zoom, retournement horizontal
Division des données	80 % entraînement, 20 % validation (via train_test_split de scikit-learn)
Format final des tenseurs	(256, 256, 1) pour les images et les masques

Tableau IV-5 : Paramètres des données et prétraitement du modèle final.

IV.6 Paramètres d'architecture du modèle U-Net final

Le modèle proposé repose sur une architecture U-Net classique, spécifiquement adaptée aux tâches de segmentation sémantique dans le domaine médical. Cette architecture est optimisée pour traiter les images échographiques en niveaux de gris, avec une configuration soigneusement choisie afin d'assurer un bon compromis entre performance et généralisation. Le tableau (Tableau IV-6) ci-dessous résume les principaux paramètres structurants du

modèle.

<i>Paramètre</i>	<i>Valeur / Description</i>
Architecture	U-Net classique avec 4 blocs d'encodage et 4 blocs de décodage
Taille d'entrée	(256, 256, 1) — images échographiques en niveaux de gris
Nombre de couches convolutives	2 par bloc (encodeur/décodeur) + 2 dans le pont central
Taille des noyaux de convolution	3 × 3 (toutes les couches convolutives)
Fonction d'activation	ReLU (pour toutes les couches internes)
Fonction d'activation finale	Sigmoïde (pour produire un masque binaire entre 0 et 1)
Régularisation	L1 = 1e-6, L2 = 1e-4 (via kernel_regularizer=l1_l2)
Fonction de perte	Dice Loss (fonction personnalisée pour la segmentation)
Optimiseur	Adam (learning_rate = 1e-4)
Métriques d'évaluation	Dice Coefficient, IoU (Intersection over Union), Binary Accuracy
Stratégie de sauvegarde	ModelCheckpoint pour sauvegarder le meilleur modèle basé sur val_dice_coef
Dropout	0.5 au niveau du pont central (bottleneck) et à la fin du bloc d'encodage le plus profond

Tableau IV-6 : Paramètres d'architecture du modèle final.

IV.7 Présentation des résultats du modèle final

Cette section présente les résultats obtenus lors de l'application du modèle de Deep Learning proposé. La figure suivante, montre que les contours des nodules sont bien reproduits, même dans les cas complexes et les formes prédites sont cohérentes confirmant

ainsi la précision morphologique du modèle. Donc, notre modèle démontre une capacité exceptionnelle à apprendre les structures fines présentes dans les images échographiques, Constituant ainsi le choix final recommandé pour la tâche de segmentation des nodules thyroïdiens dans ce projet.

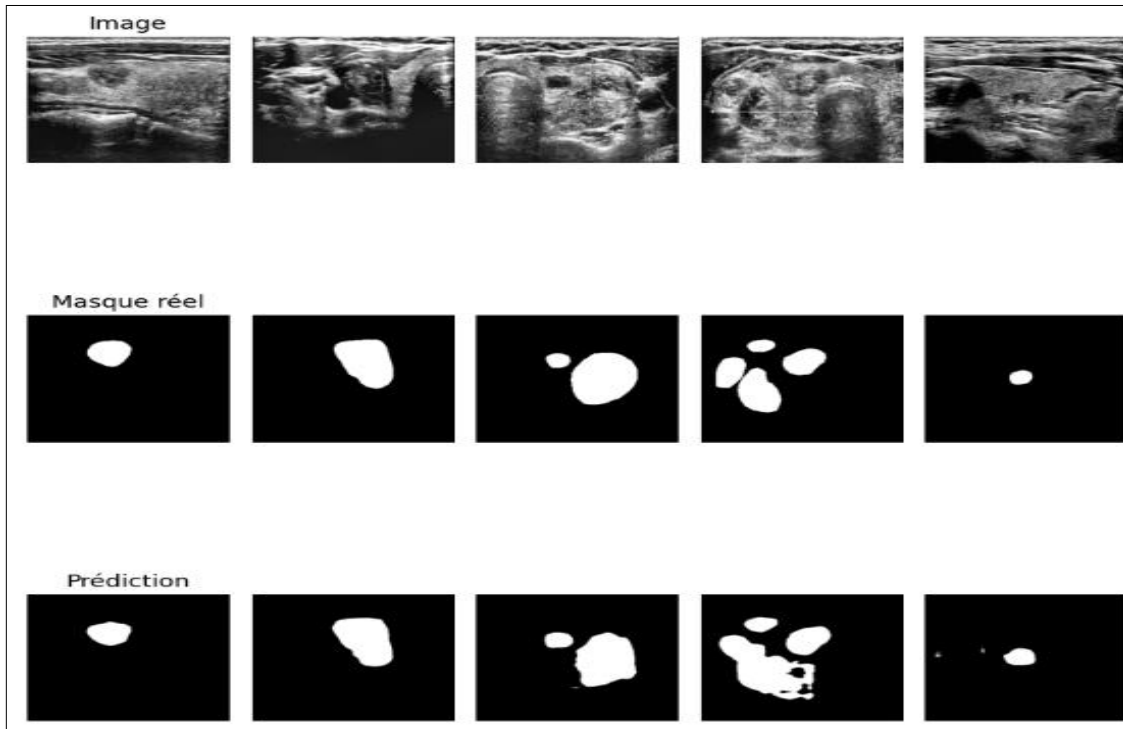
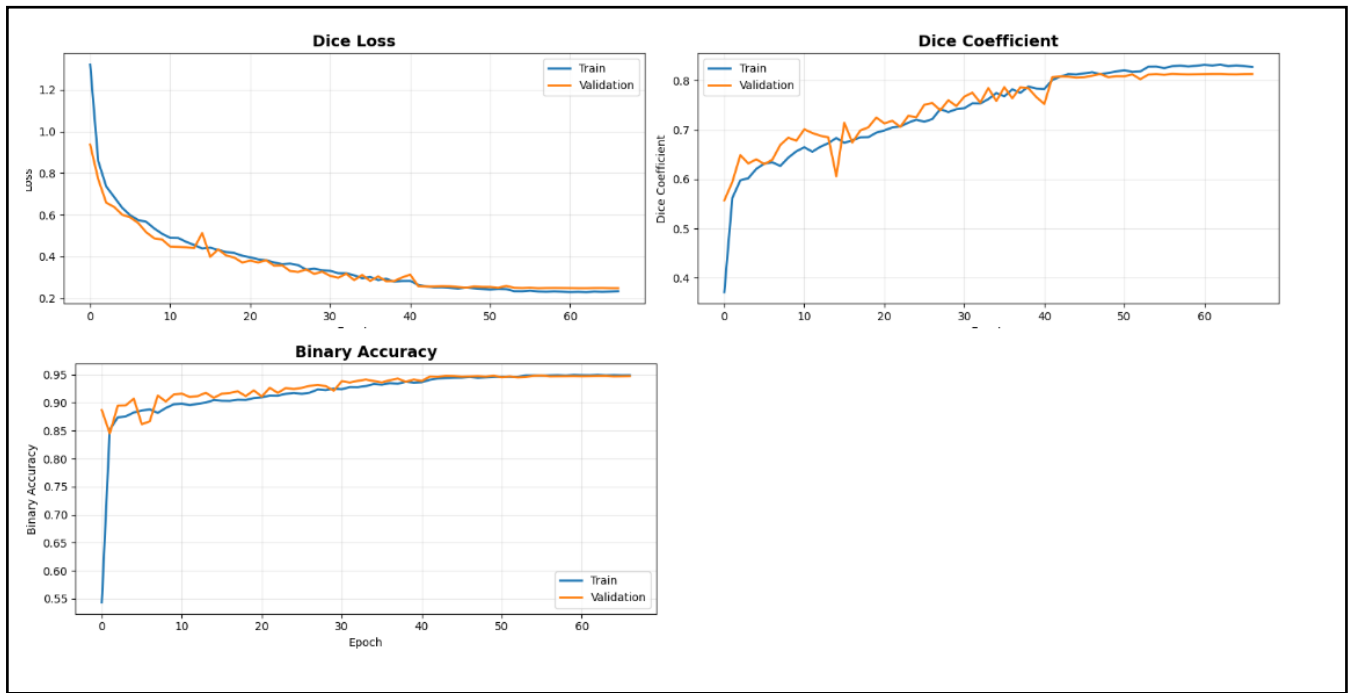


Figure IV-9 : résultats de segmentation du modèle final.

La figure suivante permet d'analyser le comportement du modèle tout au long de l'apprentissage, en fournissant un aperçu pertinent de sa performance globale ainsi que sa stabilité au fil des époques.



Test Loss: 0.2176
 Test Dice Coefficient: 0.8337
 Test Accuracy: 0.9635

Figure IV-10 : Courbes d'apprentissage du modèle final.

Figure IV-11 : Précision du modèle final.

IV.8 Comparaison des modelés

Après une évaluation des différentes architectures de segmentation, le modèle U-Net classique s'est distingué comme le meilleur compromis entre performance, stabilité, simplicité et capacité de généralisation pour la segmentation des nodules thyroïdiens. Afin de comparer objectivement les performances des différentes architectures testées, le tableau suivant synthétise les avantages et les limites observés pour chaque modèle au cours des expérimentations

Modèle	Avantages	Limites
DeepLabv3+	Bonnes métriques quantitatives (Dice, accuracy)	Masques prédits imprécis, contours absents, architecture trop lourde pour images échographiques floues
YOLOv5 + XGBoost	Pipeline intéressant pour détection + classification	Entraînement non terminé (ressources insuffisantes), surcoût computationnel trop élevé
Attention U-Net	Bon apprentissage sur les données d'entraînement	Surapprentissage à partir de l'époque 20, forme des masques erronée, sur-segmentation
U-Net classique	Excellente stabilité, courbes train/val très proches, Dice = 0.83, Accuracy > 96%	Aucun signe de surapprentissage, segmentation précise même dans les cas complexes

Tableau IV-7 : Comparaison synthétique des modèles testés

Le modèle U-Net classique a été retenu comme choix final pour la tâche de segmentation des nodules thyroïdiens, en raison de son excellent compromis entre simplicité, robustesse et précision, ce qui en fait un candidat idéal pour une utilisation dans un processus de diagnostic assisté par ordinateur basé sur l'imagerie échographique.

IV.9 Jeu de données AUITD

IV.9.1 Introduction

Le jeu de données utilisé pour la classification des nodules thyroïdiens dans ce travail est l'AUITD (Algerian Ultrasound Images Thyroid Dataset), mis à disposition publiquement sur la plateforme Kaggle par des chercheurs algériens. Il a été conçu dans le cadre d'une initiative visant à faciliter la recherche en diagnostic assisté par intelligence artificielle (IA) dans le domaine de l'endocrinologie et de l'imagerie médicale [28].

Le dataset contient des images échographiques réelles de nodules thyroïdiens, classées selon deux catégories :

- Nodules bénins.
- Nodules malins.

Chaque image est associée à une étiquette binaire (0 ou 1), permettant ainsi d'aborder un problème de classification supervisée. Ce jeu constitue une base pertinente pour un premier entraînement de modèles CNN, en particulier pour des approches de pré-diagnostic basées uniquement sur l'analyse visuelle des images.

IV.9.2 Limites du dataset AUITD pour une classification approfondie

Bien que le dataset AUITD soit classé en deux types de nodules (bénin / malin), plusieurs limites majeures empêchent l'implémentation d'une classification médicale fiable et avancée :

IV.9.2.1 Absence de masques de segmentation

L'un des verrous majeurs est l'absence totale de masques segmentés délimitant précisément les contours des nodules dans chaque image. Cela empêche toute approche hybride combinant :

- U-Net pour la segmentation automatique du nodule,
- CNN pour la classification sur la région d'intérêt (ROI).

Cette combinaison est essentielle dans un pipeline médical rigoureux, car elle permet de :

- Localiser le nodule de façon automatisée,
- Extraire des caractéristiques visuelles pertinentes (forme, bord, texture),

Et d'effectuer une classification fiable en s'appuyant sur la région segmentée, comme le fait un radiologue.

En l'absence de segmentation, la classification globale de l'image est approximative et sujette à bruit, car elle intègre des pixels non pertinents (fond, artefacts, etc.).

IV.9.2.2 Impossibilité d'utiliser le système TI-RADS

Le dataset AUITD ne contient aucune annotation clinique détaillée (forme, échogénicité, microcalcifications, contours, etc.) permettant de classer les nodules selon les critères TI-RADS (Thyroid Imaging Reporting and Data System), qui sont pourtant le standard médical actuel. Ainsi :

- Une classification multi-classe TI-RADS n'est pas faisable

Il est impossible de générer des scores de risque ou de prendre des décisions médicales fondées.

IV.9.2.3 Taille limitée du dataset et sensibilité des CNN

Les architectures CNN sont hautement dépendantes de la taille des données d'entraînement. Le dataset AUITD reste de taille modeste (quelques centaines d'images), ce qui :

- Limite la capacité de généralisation du modèle,
- Augmente le risque d'overfitting,
- Rend difficile l'entraînement de réseaux profonds sans data augmentation agressive ou transfert d'apprentissage.

Même si les étiquettes bénin/malin sont présentes, la qualité et la richesse de l'information contenue dans le dataset sont insuffisantes pour former un modèle robuste médicalement exploitable.

IV.9 CNN pour la classification binaire des nodules thyroïdiens (malin / bénin)

IV.9.3 Performance élevée sur les images médicales

Les réseaux de neurones convolutifs (CNN) sont particulièrement adaptés à l'analyse d'images, notamment les images échographiques, car ils extraient automatiquement des

caractéristiques visuelles complexes comme les bords, textures, ou formes [26]. Ces éléments sont essentiels pour différencier les nodules bénins des nodules malins, qui présentent souvent des différences très subtiles en échographie.

IV.9.4 Moins de prétraitement requis

Contrairement aux méthodes classiques comme le SVM ou les approches basées sur des règles, les CNN permettent de réduire le besoin d'ingénierie de caractéristiques manuelle, car ils apprennent les représentations optimales à partir des données brutes [26].

IV.9.5 Bonne capacité de généralisation

Les CNN, lorsqu'ils sont bien entraînés, ont montré une excellente capacité à généraliser à de nouvelles données patientes même avec des images issues de différents appareils d'échographie [29].

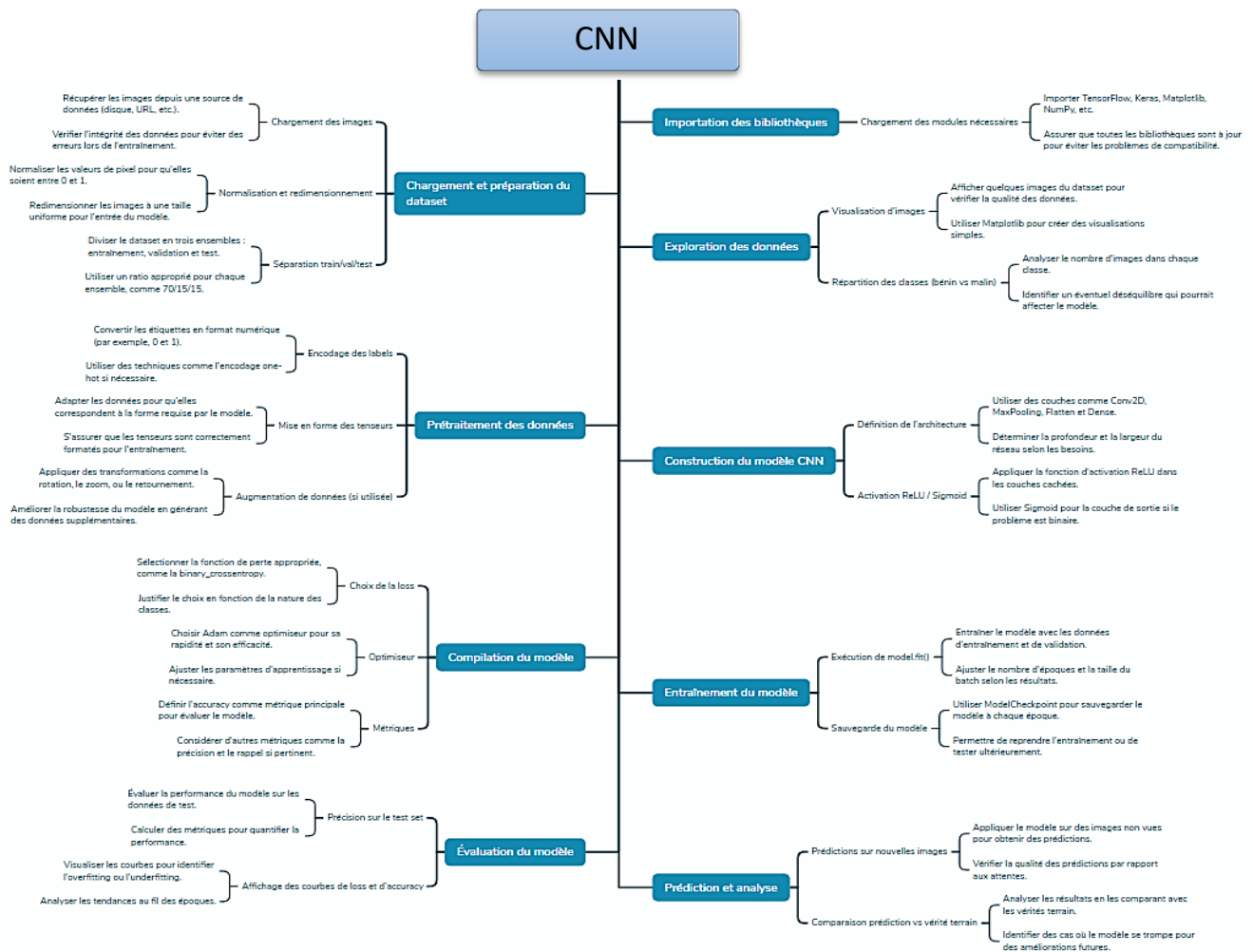
IV.9.6 Validation par la littérature scientifique

De nombreuses études récentes montrent que les CNN surpassent systématiquement les méthodes traditionnelles pour la classification des nodules thyroïdiens, avec des performances dépassant parfois celles des radiologues [26] [29].

IV.10 Étapes de réalisation du modèle de classification

La méthodologie adoptée pour la mise en œuvre du modèle de classification binaire des nodules thyroïdiens (bénin vs malin) repose sur une démarche rigoureuse et structurée. Ce processus s'appuie sur les réseaux de neurones convolutifs (CNN), largement reconnus pour leurs performances en traitement d'images médicales, afin de permettre une différenciation automatique entre les types de nodules.

L'approche est divisée en plusieurs étapes clés : depuis le chargement et la préparation du jeu de données, en passant par le prétraitement des images, la construction et la compilation du modèle CNN, jusqu'à l'entraînement, l'évaluation et l'analyse des performances du modèle.



La figure suivante illustre l'ensemble de ce pipeline de manière claire et hiérarchisée.

Figure IV-12 : Organigramme méthodologique du modèle CNN

Les principales étapes de ce processus méthodologique sont les suivantes :

- **Importation des bibliothèques**

Cette étape consiste à charger les bibliothèques indispensables au traitement des images, à la création du modèle de classification par Deep Learning, à son entraînement, son évaluation ainsi qu'à la visualisation des résultats. Les principaux modules utilisés sont TensorFlow, Keras, NumPy, Matplotlib et scikit-learn.

- **Chargement et prétraitement des données**

Les images issues de la base de données sont lues, redimensionnées à une taille standard (ex. 224×224), et normalisées (valeurs de pixels entre 0 et 1). Des générateurs d'images (ImageDataGenerator) permettent aussi l'augmentation des données pour enrichir le jeu d'entraînement (flip, rotation, zoom), tout en assurant une répartition adéquate entre les ensembles d'entraînement, validation et test.

- **Visualisation du prétraitement**

Cette phase permet de visualiser quelques échantillons d'images brutes et prétraitées afin de vérifier la qualité des transformations appliquées. Des graphiques sont également générés pour analyser la répartition des classes (bénin vs malin) et détecter d'éventuels déséquilibres dans le dataset.

- **Construction du modèle CNN**

Un modèle de réseau de neurones convolutif (CNN) est construit à l'aide de Keras. Il comprend plusieurs couches : convolutives, de pooling, de flattening et de couches denses. Les fonctions d'activation ReLU et Sigmoid sont utilisées respectivement pour les couches cachées et la couche de sortie, cette dernière étant adaptée au problème de classification binaire.

- **Compilation du modèle**

Le modèle est compilé avec l'optimiseur Adam, la fonction de perte binary_crossentropy, et la métrique accuracy pour évaluer les performances. Cette étape prépare le modèle pour l'entraînement.

- **Entraînement du modèle**

Le modèle est entraîné sur les données d'apprentissage avec validation sur un ensemble dédié. Plusieurs **callbacks** sont intégrés :

- **EarlyStopping** pour arrêter l'entraînement si les performances stagnent,
- **ReduceLROnPlateau** pour ajuster dynamiquement le taux d'apprentissage,
- **ModelCheckpoint** pour sauvegarder automatiquement le meilleur modèle.

- **Évaluation du modèle**

Le modèle entraîné est évalué sur l'ensemble de test. Les métriques de performance telles que la précision, la perte et d'autres mesures (comme la matrice de confusion, le rappel et la F1-score) sont calculées à l'aide des fonctions de `sklearn.metrics`. Cela permet d'estimer la capacité du modèle à généraliser sur de nouvelles données.

- **Sauvegarde du modèle**

Le modèle entraîné est sauvegardé au format `.h5`, ce qui permet de le recharger ultérieurement pour l'inférence ou pour un déploiement dans une application web ou mobile.

- **Visualisation des résultats**

Les courbes de perte et d'exactitude (loss/accuracy) pour les phases d'apprentissage et de validation sont tracées afin d'analyser le comportement du modèle au fil des époques. En complément, les prédictions sur des images test sont comparées aux vraies classes à l'aide de la matrice de confusion et du rapport de classification pour une évaluation qualitative et quantitative.

IV.11 Paramètres du modèle CNN utilisée

Afin d'assurer une reproduction rigoureuse de l'expérimentation et de mieux comprendre les choix techniques effectués, le tableau ci-dessous présente l'ensemble des paramètres utilisés pour l'entraînement du modèle de classification des nodules thyroïdiens. Ces paramètres ont été définis en fonction des bonnes pratiques en deep learning, ainsi que des spécificités des données médicales traitées (images échographiques). Le modèle repose sur une architecture CNN adaptée à la classification binaire, et intègre des techniques d'optimisation et de régularisation visant à maximiser les performances tout en minimisant le risque de surapprentissage.

<i>Paramètre</i>	<i>Valeur</i>
Type de modèle	CNN (Convolutional Neural Network)

Taille des images en entrée	224×224×3(images échographiques en RGB ou converties)
Nombre de classes	2 (bénin, malin)
Fonction d'activation (couches cachées)	ReLU
Fonction d'activation (couche de sortie)	Sigmoid (classification binaire)
Fonction de perte	binary_crossentropy
Optimiseur	Adam
Taux d'apprentissage (learning rate)	0.001
Métriques utilisées	Accuracy
Taille du batch	32
Nombre d'époques	50 (early stopping dans l'époque 21)
Callbacks	EarlyStopping,ModelCheckpoint, ReduceLROnPlateau
Normalisation des images	Oui

Techniques d'augmentation	Rotation, Zoom, HorizontalFlip
Prétraitement avancé	Égalisation d'histogramme (CLAHE)
Format de sauvegarde du modèle	.h5

Tableau IV-8 : Paramètres d'entraînement du modèle de classification par CNN

IV.12 Paramètres d'architecture du modèle CNN

Le modèle de classification utilisé repose sur une architecture de réseau de neurones convolutifs (CNN), spécialement conçue pour traiter les images échographiques et distinguer les nodules bénins des nodules malins. Cette architecture est relativement simple mais efficace, avec une structure hiérarchique de couches convolutionnelles et de pooling permettant d'extraire des caractéristiques discriminantes à partir des images. Elle se termine par des couches denses pour la prise de décision. Le tableau ci-dessous présente les principaux paramètres structurants de cette architecture.

Paramètre	Valeur
Nombre total de couches	10 (3 Conv2D, 3 MaxPooling2D, 1 Flatten, 2 Dense, 1 Dropout)
Nombre de filtres convolutionnels	32 → 64 → 128 (augmentation progressive par bloc)
Taille des noyaux de convolution	(3 × 3), padding same
Taille des fenêtres de pooling	(2 × 2) pour chaque couche de MaxPooling2D

Couche de régularisation	Dropout (taux = 0.5 après couche dense)
Couche de sortie	Dense(1) avec activation Sigmoid (pour classification binaire)
Nombre total de paramètres entraînables	Environ 1,6 million (selon le résumé du modèle)

Tableau IV-9 : Paramètres d'architecture du modèle CNN

IV.13 Analyse des performances du modèle de classification

L'évaluation du modèle CNN pour la classification des nodules thyroïdiens a été réalisée sur un ensemble de test indépendant, à l'aide de plusieurs outils d'analyse, courbes d'apprentissage, matrice de confusion et visualisation des erreurs de prédiction.

IV.13.1 Courbes d'apprentissage

Les graphiques représentant l'évolution de la loss et de l'accuracy au fil des époques montrent une tendance à la stagnation après quelques itérations. La courbe d'accuracy atteint un plafond relativement bas, et la perte ne diminue pas significativement, indiquant une capacité d'apprentissage limitée du modèle. Cela peut être le signe :

- Un underfitting (le modèle n'a pas appris suffisamment de motifs discriminants),
- Un volume de données insuffisant pour un apprentissage profond efficace.

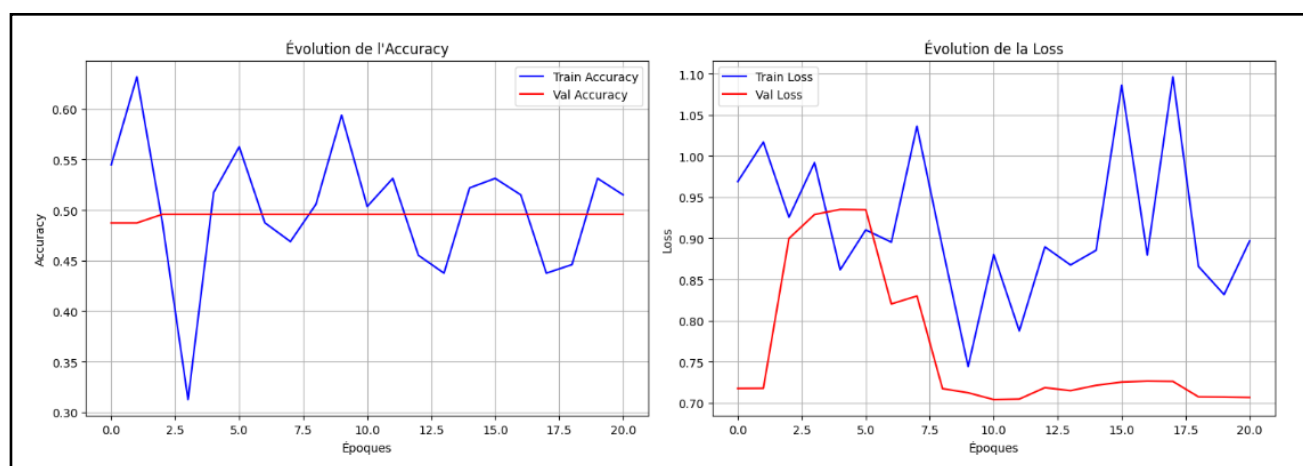


Figure IV-13 : Courbes d'apprentissage du modèle CNN.

IV.13.2 Matrice de confusion

La matrice de confusion obtenue met en évidence une confusion importante entre les deux classes (bénin et malin). On observe notamment :

- Un grand nombre de faux positifs (nodules bénins classés à tort comme malins),
- Une précision et un rappel faibles pour la classe bénigne, ce qui déséquilibre le modèle en faveur de la classe "malin". Cela confirme que le modèle a tendance à sur-prédire la classe majoritaire ou la plus marquée dans les exemples d'entraînement, ce qui affecte sa fiabilité clinique.

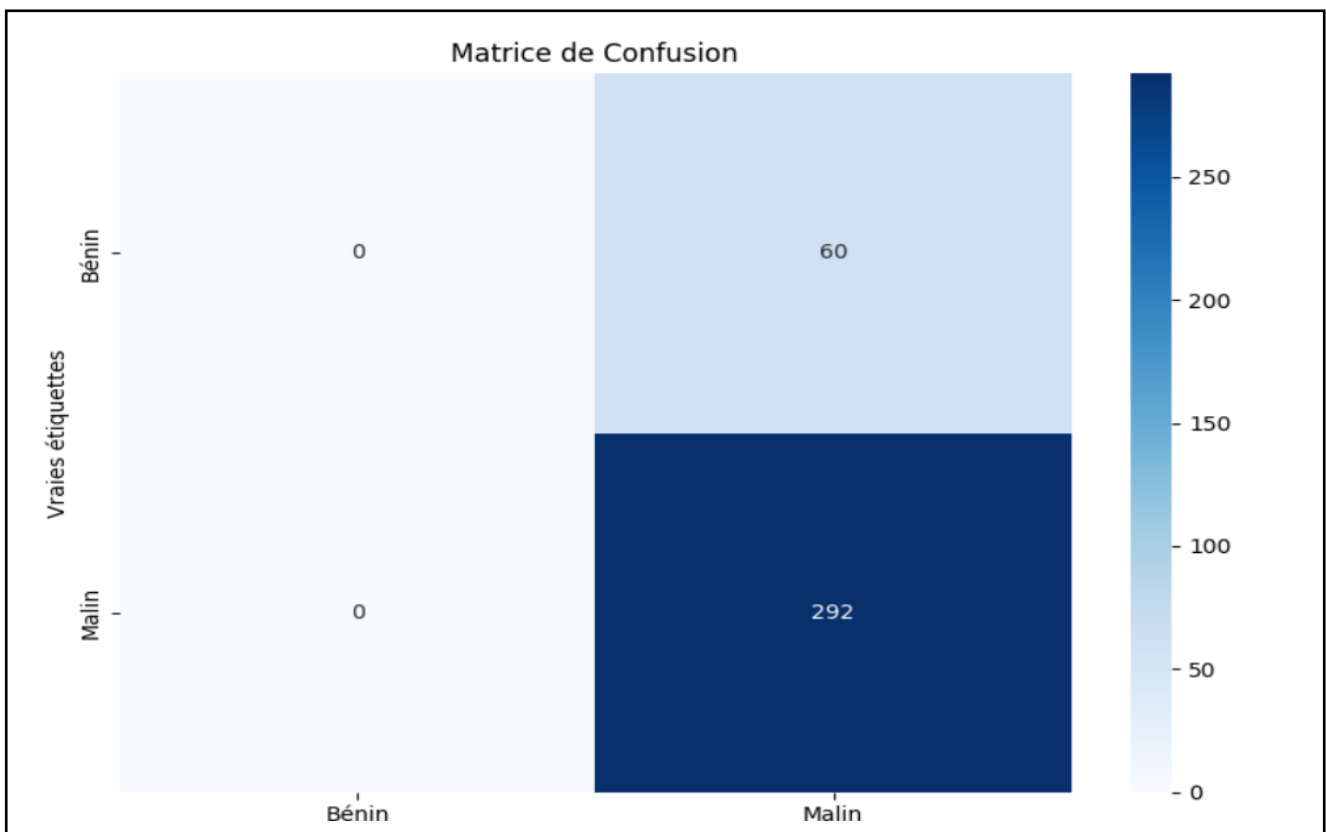


Figure IV-14 : Matrice de confusion pour le modèle CNN

IV.13.3 Visualisation des erreurs

L'image ci-dessous illustre plusieurs exemples de faux positifs où le modèle a prédit la présence d'un nodule malin alors que le diagnostic réel était bénin (pour tous les cas bénins).

Ces erreurs traduisent l'incapacité du modèle à bien distinguer les motifs visuels caractéristiques, souvent subtils dans les images échographiques en niveaux de gris.

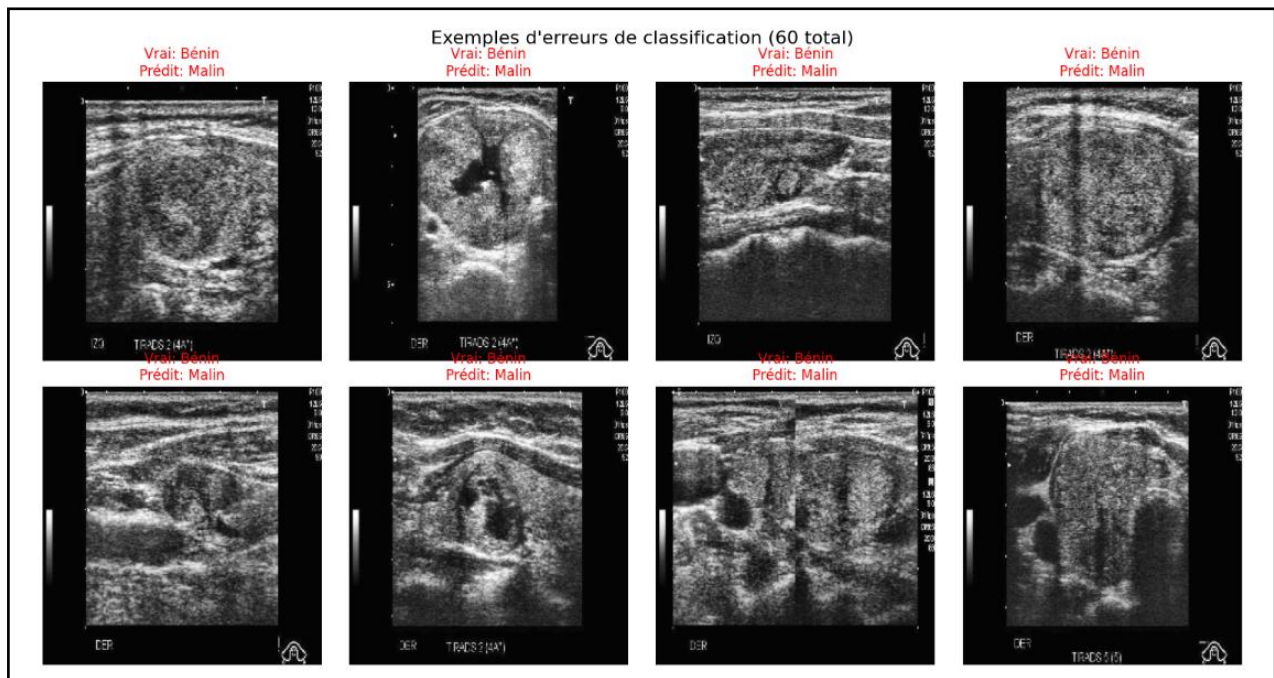


Figure IV-15 : résultats de classification du modèle.

IV.13.4 Discussion et limitations

Les résultats obtenus soulignent plusieurs limitations majeures :

- Le jeu de données est trop restreint pour entraîner efficacement un modèle CNN, qui nécessite un grand volume de données pour bien apprendre les caractéristiques discriminantes.
- Le déséquilibre entre classes dans la base de données accentue les erreurs de classification, en particulier les faux positifs.
- Manque de caractéristiques pour faire la classification (type, le rapport longueur/largeur, composition (liquide, solide, kelsyque)).

IV.14 Conception d'une Application Web

L'interface utilisateur de l'application web a été conçue pour être à la fois intuitive et fonctionnelle, permettant aux professionnels de santé ou chercheurs de visualiser de manière claire les résultats des traitements par Deep Learning appliqués aux images échographiques thyroïdiennes.

Trois modes d'utilisation sont proposés :

- Segmentation uniquement,
- Classification uniquement,
- Segmentation + Classification combinées.

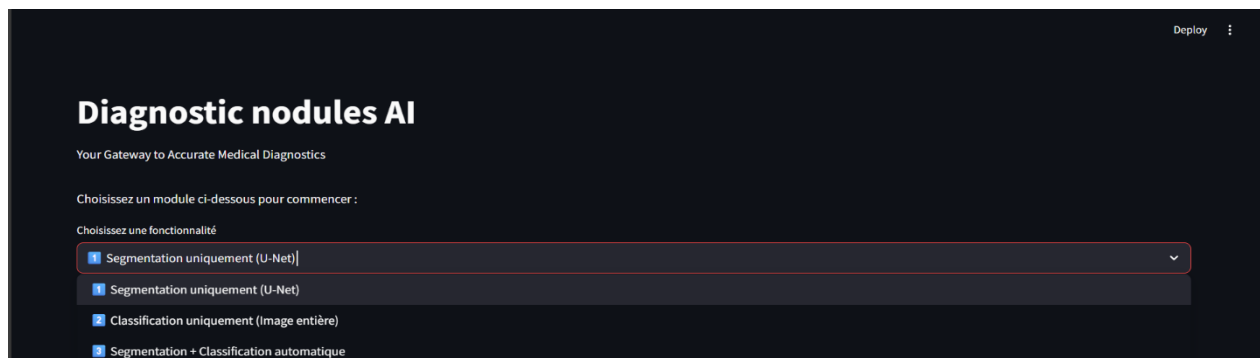


Figure IV-16 : L'interface d'application

Dans le mode Segmentation, l'interface permet de charger une image échographique, qui sera ensuite traitée automatiquement par le modèle (U-Net pour la segmentation). Les résultats sont ensuite affichés sous forme de trois volets visuels :

- **Image originale** : l'image échographique telle qu'importée par l'utilisateur, sans modification.
- **Masque prédit** : ce masque est généré par le modèle U-Net, et montre la région du nodule identifiée automatiquement.
- **Superposition masque/image** : une version de l'image originale sur laquelle le masque prédit est superposée en couleur (en gris semi-transparent), afin de visualiser directement l'emplacement estimé du nodule sur l'image échographique.



Figure IV-17 : Résultat de segmentation d'une image du nodule via l'interface

Dans le mode classification, l'interface affiche l'image traitée par le modèle (CNN) en sortie, accompagnée au-dessus de celle-ci de la classe prédite par le modèle ainsi que du pourcentage de confiance associé à cette prédiction. Cette présentation permet à l'utilisateur de visualiser clairement le résultat du traitement, en identifiant non seulement la catégorie du nodule thyroïdien (bénin ou malin), mais aussi le degré de certitude du modèle quant à cette classification. Ainsi, cette information facilite la prise de décision médicale en fournissant une indication explicite de la fiabilité de la prédiction.



Figure IV-18 : Résultat de classification d'une image du nodule via l'interface

Dans le mode segmentation + classification, l'interface affiche côte à côte trois images : l'image originale, ainsi qu'une image de superposition du masque de segmentation sur l'image traitée, l'image traitée contenant le nodule extrait. Au-dessus de ces trois images, sont affichés les résultats de la classification, comprenant la classe prédite par le modèle (malin ou bénin) ainsi que le pourcentage de confiance associé à cette prédiction. Cette présentation permet à l'utilisateur d'avoir une vision complète du processus, en visualisant à la fois la localisation précise du nodule grâce à la segmentation, ainsi que l'interprétation clinique apportée par la classification avec un indicateur clair de la fiabilité du diagnostic.

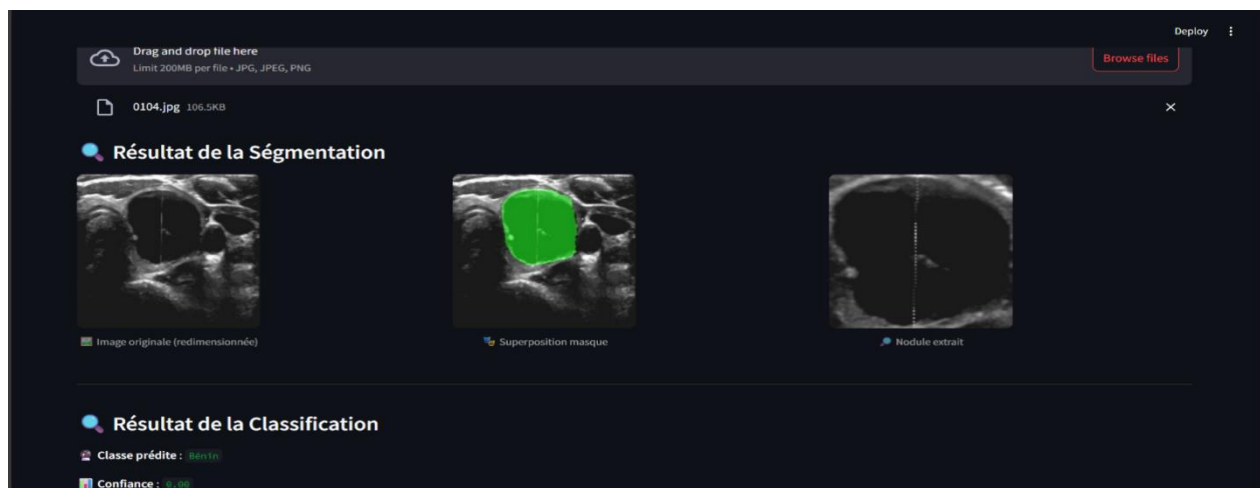


Figure IV-19 : Résultat de segmentation + classification d'une image du nodule via l'interface

IV.14.1 Remarques

Les résultats de classification obtenus ne sont pas entièrement fiables, principalement en raison des limitations du modèle de classification utilisé. En effet, la précision du diagnostic dépend fortement de la qualité et de la robustesse du modèle, qui peut parfois se tromper ou manquer de confiance sur certains cas.

Concernant le mode segmentation + classification, la qualité des résultats est satisfaisante pour la partie segmentation, ce qui permet une bonne extraction du nodule sur l'image traitée. En revanche, la fiabilité des résultats globaux dépend essentiellement de la performance du

modèle de classification, qui reste perfectible. Ainsi, même si la segmentation est correcte, la classification peut ne pas être juste, ce qui impacte la pertinence finale de l'interprétation clinique.

IV.15 Limites de la classification dans notre cas spécifique

Bien que les CNN aient démontré leur efficacité pour la classification binaire des nodules thyroïdiens dans de nombreuses études [26] [29], cette approche ne peut pas être appliquée de manière fiable dans notre cas. En effet, la base de données utilisée, AUITD, présente plusieurs limitations majeures qui compromettent l'entraînement d'un modèle de classification robuste. D'abord, le nombre total d'images est insuffisant pour permettre à un CNN d'apprendre des représentations discriminantes solides, ce qui augmente fortement le risque de surapprentissage. Ensuite, la base ne contient pas de masques de segmentation, ce qui empêche toute localisation précise des nodules et rend difficile l'extraction de régions d'intérêt pertinentes. Enfin, les métadonnées cliniques nécessaires à la classification (telles que les résultats histopathologiques, les caractéristiques échographiques clés ou le diagnostic final bénin/malin) sont absentes ou incomplètes. Ces lacunes rendent la classification automatique peu fiable, voire inapplicable dans le contexte de notre étude.

IV.16 Conclusion

Ce chapitre a permis de mettre en œuvre et d'évaluer expérimentalement le modèle de segmentation basé sur l'architecture U-Net, appliqué aux images échographiques de nodules thyroïdiens. Les différentes séries de tests ont démontré que l'entraînement sur 75 époques, couplé à des techniques d'amélioration telles que la régularisation, l'ajustement dynamique du taux d'apprentissage et l'early stopping, conduit à une convergence stable et à une amélioration notable des performances. Les métriques quantitatives (Dice, IoU, accuracy) ainsi que les résultats visuels obtenus montrent que le modèle proposé est capable de générer des segmentations précises, cohérentes et adaptées aux particularités des images médicales.

Enfin, la comparaison avec d'autres architectures comme DeepLabv3+, Attention U-Net ou encore YOLOv5, a confirmé la supériorité du modèle U-Net dans ce contexte spécifique, tant en termes de qualité des masques produits que de stabilité d'apprentissage.

Ainsi, ces résultats valident l'efficacité du modèle U-Net pour la segmentation automatique des nodules thyroïdiens, et posent les bases d'une intégration fiable dans une chaîne de diagnostic assisté par intelligence artificielle.

CONCLUSION GENERALE

Conclusion Générale

Ce mémoire s'est attaché à développer un système automatique de segmentation des nodules thyroïdiens à partir d'images échographiques, en exploitant les capacités des réseaux de neurones convolutifs profonds et plus particulièrement l'architecture U-Net.

À travers une étude structurée, nous avons d'abord rappelé l'importance du dépistage précoce des nodules dans le contexte médical, avant de présenter les concepts fondamentaux du Deep Learning appliqué à l'imagerie médicale.

Dans la phase expérimentale, plusieurs architectures ont été testées et comparées. L'évaluation a porté sur des métriques standard telles que le Dice coefficient, l'Intersection over Union (IoU), et l'accuracy, ainsi que sur la qualité visuelle des masques segmentés.

À l'issue de ces expérimentations, le modèle U-Net classique, entraîné sur 75 époques avec une régularisation L1/L2, un apprentissage adaptatif (ReduceLROnPlateau) et un arrêt anticipé (EarlyStopping), s'est imposé comme le modèle final retenu. Il a démontré une excellente stabilité d'apprentissage, une capacité de généralisation élevée, et des résultats très satisfaisants sur les données de test, atteignant :

- Dice coefficient de 0.8337
- Accuracy de 0.9635

Ce modèle a également surclassé d'autres approches testées, telles que DeepLabv3+, Attention U-Net et YOLOv5, tant en précision quantitative qu'en cohérence visuelle.

Ainsi, l'architecture U-Net classique optimisée s'est révélée être une solution efficace et fiable pour la segmentation automatique des nodules thyroïdiens, apportant une réelle valeur ajoutée dans le contexte du diagnostic assisté par intelligence artificielle.

Perspectives

Les résultats obtenus dans ce mémoire constituent une base solide pour la mise en œuvre d'un système complet d'aide au diagnostic. Plusieurs perspectives de prolongement peuvent être envisagées :

Classification bénin/malin à partir des régions segmentées :

Une suite logique à la segmentation est l'extraction automatique du nodule afin d'entraîner un modèle de classification binaire (bénin vs malin). Cette approche permettrait d'assister les médecins dans l'estimation du risque tumoral. Des premiers tests montrent que cette tâche est réalisable, mais reste conditionnée par la qualité des masques segmentés.

Difficultés liées aux jeux de données disponibles :

L'une des principales limitations rencontrées concerne la disponibilité de jeux de données adaptés. La majorité des datasets publics ne fournissent qu'un diagnostic binaire (0/1) sans informations cliniques détaillées (forme, échogénicité, bord, microcalcifications, etc.), ce qui restreint les capacités de généralisation des modèles et freine le développement de systèmes plus complets.

Vers une classification multi-classes selon le score TI-RADS :

Pour aller plus loin, il serait pertinent de développer un système capable de prédire automatiquement une classe TI-RADS (Thyroid Imaging Reporting and Data System), ce qui reflète la pratique médicale réelle. Toutefois, cela nécessite un dataset riche en annotations structurées, avec des milliers d'images étiquetées par des radiologues selon les critères TI-RADS. À l'heure actuelle, ces jeux de données sont rares, voire inexistants publiquement.

Une piste intéressante serait donc de collaborer avec des institutions médicales pour construire un dataset plus complet, intégrant à la fois les images, les masques, les scores TI-RADS et les diagnostics histopathologiques.

BIBLIOGRAPHIE

Bibliographie

- [1] : Boubekri Nassima (2020). Chapitre III: La physiologie de la glande Thyroïde. Faculté de Médecine, université de Mostaganem (Algérie).
- [2]: Williams, P. L. (2019). Gray Anatomy for Students (4th ed.). Elsevier Health Sciences.
- [3]: Hall, J. E. (2021). Guyton and Hall Textbook of Medical Physiology (14th ed)
- [4]: Ajilisa, A., Sreedhar, C., & Sundararajan, V. (2023). Artificial intelligence for thyroid nodule diagnosis from ultrasound images
- [5]: Ma, J., & Zhang, Y. (2022). Deep learning for differentiating malignant and benign thyroid nodules on ultrasound: A review. *Frontiers in Oncology*
- [6]: Vahdati, A., Tareke, A. A., et al. (2024). Ultrasound-based AI models for thyroid disorder detection. *Bioengineering*,
- [7]: Tessler, F. N., Middleton, W. D., Grant, E. G., et al. (2017). ACR TI-RADS: White Paper of the ACR TI-RADS Committee. *Journal of the American College of Radiology (JACR)*
- [8]: Haugen, B. R., et al. (2016). 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*
- [9]: Cibas, E. S., & Ali, S. Z. (2017). The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*,
- [10]: Bongiovanni, M., et al. (2012). The Bethesda System for Reporting Thyroid Cytopathology: A Meta-analysis. *Acta Cytologica*
- [11]: Baloch, Z. W., & Livolsi, V. A. (2019). Diagnostic pathology of thyroid gland. In *Diagnostic Pathology: Endocrine* (3rd ed., pp. 239–272). Elsevier.
- [12]: Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.)
- [13]: Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*
- [14]: Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [15]: 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*
- [16]: Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*
- [17]: Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*
- [18]: Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*
- [19]: Liu, F., Zhou, Z., Jang, H., Samsonov, A., Zhao, G., Kijowski, R., & Li, X. (2017). Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. *Magnetic Resonance in Medicine*

- [20]: Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems
- [21]: TN3K Dataset. (s.d.). Thyroid Nodule 3000 Dataset (TN3K) [Data set]. Kaggle.
- [22]: Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention
- [23]: Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Medical Imaging,
- [24]: Song, J., Choi, H. I., Park, H. K., & Baek, J. H. (2019). Ultrasound image analysis using deep learning for the diagnosis of thyroid nodules. Medicine,
- [25]: Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Deep convolutional neural network for the automated diagnosis of thyroid cancer using ultrasound images.
- [26]: LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning.
- [27]: Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. In International Conference on Artificial Neural Networks.
- [28]: AUITD Dataset. (2022). Artificially Unbiased International Thyroid Dataset (AUITD). Kaggle.
- [29]: 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