

Faculté des Sciences
Département
d'Informatique

Domaine **:** Mathématique et Informatique
Filière **:** Informatique
Spécialité **:** Systèmes d'Information
N° de l'Arrêté d'habilitation de la spécialité : arrêté n °872 du 26/07/2016

<u>Mémoire de fin d'études en vu de l'obtention du</u>
<u>Diplôme de Licence Académique</u>

Thème

Détection et Localisation des Microcalcifications Mammographiques par Apprentissage Profond basé sur un Modèle CNN et Transfert de Connaissances

Realisé par **:**
IGHIL GUITOUN sabrina
ACHEMOUN Nihad
HAMICHE Loubna Imane
DOUMI racim houssam Edaine

Soutenu le …./…./ ….. Devant le jury composé de
Encadreur : **Dr. TOUAZI Fayçal**
Examinateur : **Dr. BENZENATI Tayeb**

Remerciements

*Nous remercions avant tout **Allah**, le Tout-Puissant, de nous avoir accordé la santé, la volonté et la patience nécessaires pour mener à bien ce travail.*

*Nous exprimons notre profonde gratitude à notre encadreur **M. TOUAZI**, pour son accompagnement, ses conseils avisés, sa disponibilité et ses précieuses orientations tout au long de la réalisation de ce mémoire.*

Nos sincères remerciements s'adressent également aux membres du jury qui ont accepté d'évaluer notre travail.

Nous remercions chaleureusement tous nos enseignants du Département d'Informatique pour les connaissances et les compétences qu'ils nous ont transmises durant notre parcours universitaire.

Enfin, nous adressons nos plus vifs remerciements à nos familles, amis et collègues pour leur soutien moral, leurs encouragements et leur confiance tout au long de nos études.

Racim, Nihad, Loubna, Sabrina

Dédicases

Nous dédions ce modeste travail :

À nos chers parents, pour leurs sacrifices, leur amour inconditionnel, leur soutien et leurs encouragements permanents.

À nos frères et sœurs, pour leur affection et leur présence à nos côtés.

À nos enseignants, qui ont contribué à notre formation et à notre réussite.

À tous nos amis et collègues, qui ont partagé avec nous les moments marquants de notre parcours universitaire.

À toute personne ayant contribué de près ou de loin à l'aboutissement de ce travail.

Résumé :

Dans ce travail, nous avons proposé une approche basée sur l'apprentissage profond pour la détection et la localisation des masses mammaires dans les images mammographiques. Notre méthode repose sur l'utilisation du modèle YOLOv11 appliqué au dataset VinDr-Mammo.

Deux approches ont été étudiées : un entraînement sans transfer learning et un entraînement utilisant des poids pré-entraînés sur le dataset COCO. Les expériences ont été réalisées sur la plateforme Kaggle en utilisant le framework PyTorch.

Les performances des modèles ont été évaluées à l'aide des métriques mAP50 et mAP@50-95. Les résultats obtenus montrent l'efficacité du modèle YOLOv11 pour la détection des anomalies mammaires, tout en mettant en évidence les limites liées à la taille et à la complexité des données médicales.

Mots clés : Mammographie, YOLOv11, Deep Learning, Transfer Learning, VinDr-Mammo, Détection des masses.

Abstract:

In this work, we proposed a deep learning-based approach for the detection and localization of breast masses in mammographic images. Our method relies on the YOLOv11 model applied to the VinDr-Mammo dataset.Two approaches were investigated: training without transfer learning and training using pretrained weights from the COCO dataset. The experiments were conducted on the Kaggle platform using the PyTorch framework.

The performance of the models was evaluated using the mAP50 and mAP@50-95 metrics. The obtained results demonstrate the effectiveness of the YOLOv11 model for detecting breast abnormalities while highlighting the limitations related to the size and complexity of medical data.

Keywords: Mammography, YOLOv11, Deep Learning, Transfer Learning, VinDr-Mammo, Breast Mass Detection.

Liste des abréviations

- **BI-RADS** : Breast Imaging Reporting and Data System
- **CNN** : Convolutional Neural Network
- **YOLO** : You Only Look Once
- **ML** : Machine Learning
- **DL** : Deep Learning
- **ROI** : Region Of Interest
- **MN** : Mammographie Numérique
- **FC** : Fully Connected
- **SVM** : Support Vector Machine
- **KNN** : K-Nearest Neighbors
- **GPU** : Graphics Processing Unit
- **IoU** : Intersection over Union
- **mAP** : Mean Average Precision
- **COCO** : Common Objects in Context
- **CAD** : Computer-Aided Diagnosis
- **AI** : Artificial Intelligence
- **ImageNet** : Image Database for Visual Recognition

Table des matières

Détection et Localisation des Microcalcifications Mammographiques par Apprentissage Profond basé sur un Modèle CNN et Transfert de Connaissances 1

Résumé : 4

Liste des abréviations 5

Table des figures 7

Introduction Générale : 8

1. Fondements médicaux et techniques de la détection des microcalcifications 10

1.1 Introduction : 10

1.2 L'imagerie mammaire : 10

1.2.1 Mammographie : 10

Dataset 11

Vindr-Mammo 11

Nombre d'images 11

20000 11

Type 11

Mammographies FFDM 11

Annotation 11

Bounding box 11

Classes 11

Masse bénigne/maligne 11

Split 11

Train 70% /Val 10% /Test 20% 11

Table 1 represente la description du dataset Vindr-Mammo 11

1.3 Les anomalies mammaires : 11

1.4 Intelligence artificielle et Machine Learning : 11

1.5 Les calcifications mammaires : 12

1.6 Classification BI-RADS : 13

1.7 Classification automatique des images : 14

1.7.1 Définition : 14

1.7.2 Objectif : 14

1.8 L'apprentissage automatique : 14

1.9 L'apprentissage profond "Deep Learning" : 14

1.10 Les réseaux de neurones convolutifs : 15

1.10.1 La convolution : 15

1.10.2 Le Pooling : 16

1.10.3 Les couches entièrement connectées : 16

1.11 Transfert des connaissances « Tranfert learning » : 17

1.12 L'architecture YOLO : 17

1.13 Métriques d'évaluation : 18

Conclusion : 19

2.1 Introduction : 15

2.2 Datasets utilisés : 15

2.2.1 Dataset CBIS-DDSM : 15

2.2.2 Dataset VinDr-Mammo :....16
2.3 Prétraitement des données : 17
2.4 Environnement de travail : 17
2.4.1 Kaggle : 17
2.4.2 PyTorch : 18
2.4.3 Ultralytics YOLO : 18
2.5 Architecture YOLOv11 : 19
- Backbone : 19
- Neck : 19
- Head : 19

2.6 Transfer Learning : 20
2.7 Dataset COCO : 20
2.8 Paramètres d'entraînement :....21
2.9 Métriques d'évaluation : 21
2.9.1 Intersection over Union (IoU) :.... 21
2.9.2 Dice Coefficient : 22
2.9.3 mAP50 : 23
2.9.4 mAP50-95 : 23
2.10 Les approches proposées : 23
2.10.1 Approche sans Transfer Learning : 23
2.10.2 Approche avec Transfer Learning :.... 23
Conclusion :....23
3.2 Comparaison des modèles utilisés : 26
3.3 Analyse des résultats :....26
3.5 Limites de l'étude :....27
Conclusion :....30

Table des figures

- Figure 1 : Mammographie numérique (MN) et les clichés associés
- Figure 2 : Différentes formes des microcalcifications
- Figure 3 : Mammographie montrant un foyer de microcalcifications et agrandissement
- Figure 4 : Différence entre l'apprentissage automatique et l'apprentissage profond
- Figure 5 : Architecture d'un réseau de neurones convolutifs (CNN)
- Figure 6 : Représentation de la convolution dans une image
- Figure 7 : Réduction de la taille de la carte de caractéristiques (Feature Map)
- Figure 8 : Exemple d'opération de Pooling (Max Pooling 2×2)
- Figure 9 : Couche entièrement connectée (Fully Connected Layer)
- Figure10 :schema de yolo
- Figure11:Exemple de dataset CBIS-DDSM
- Figure12 :exemple de dataset VinDR Mammo
- Figure13 Interface de la plateforme Kaggle utilisée pour l'entraînement des modèles de Deep Learning .
- Figure 14 :exemple de dataset COCO
- Figure X.X :illustration de l'intersection over union (loU)
- Figure 15:loU evaluation example

- Figure 16:YOLO training metrices
- Figure 17:YOLO detection examples

Liste des Tables :

- **Table 1 represente la description du dataset Vindr-Mammo**
- **Tableau 2 :Lexique BI-RADS**

- **Tableau 3 : hyperparametres lors de l'entrainemen**

Introduction Générale :

Le cancer du sein est l'une des principales préoccupations de santé publique dans le monde, touchant des millions de femmes chaque année. Il représente un défi majeur en termes de diagnostic précoce et de traitement efficace. Dans le cadre de notre mémoire de fin d'études, nous nous sommes intéressés à la classification des images mammographiques, une approche essentielle pour le dépistage précoce de cette maladie dévastatrice.

Le cancer du sein est le cancer le plus fréquent chez les femmes, avec des statistiques alarmantes qui indiquent une augmentation continue de son incidence. Selon l'Organisation mondiale de la santé (OMS), il est estimé qu'environ 2,3 millions de nouveaux cas de cancer du sein, sont diagnostiqués chaque année, et il est responsable de plus de 600 000 décès dans le monde. Ces chiffres soulignent l'urgence de développer des méthodes avancées et précises pour la détection précoce de cette maladie.

La mammographie, qui consiste en l'imagerie radiographique du sein, est l'un des outils les plus couramment utilisés dans le dépistage précoce du cancer du sein. Elle permet de détecter des anomalies potentiellement malignes, avant même l'apparition de symptômes cliniques, offrant ainsi de meilleures chances de guérison et de traitement efficace. Toutefois, l'interprétation des images mammographiques peut être complexe et subjective, ce qui souligne le besoin de méthodes automatisées pour faciliter le processus de diagnostic.

Dans notre mémoire de fin d'études, nous avons donc exploré l'utilisation des réseaux de neurones convolutionnels (CNN), pour la classification des images mammographiques. Les CNN sont des modèles d'apprentissage profond capables d'apprendre des caractéristiques visuelles complexes à partir des images, ce qui en fait des outils puissants pour la classification automatique.

Notre étude se concentre sur deux aspects clés de la classification des images mammographiques : la classification multi-classe basée sur le système BIRADS, qui permet de caractériser la probabilité de malignité des lésions détectées, et la classification binaire visant à différencier les images normales des images présentant des masses suspectes.

En résumé, notre mémoire de fin d'études propose une approche novatrice, basée sur les CNN pour la classification des images mammographiques. Nous visons à contribuer à la recherche en fournissant des méthodes avancées et automatisées, pour le dépistage précoce du cancer du sein, dans l'espoir de réduire l'impact de cette maladie sur la santé des femmes, et d'améliorer leur prise en charge médicale.

CHAPITRE 01

Fondements médicaux et techniques de la détection des microcalcifications

1. Fondements médicaux et techniques de la détection des microcalcifications

1.1 Introduction :

Ce chapitre présente le contexte médical et les bases théoriques nécessaires à notre étude sur la détection des microcalcifications mammaires. Nous introduisons l'imagerie mammaire et la mammographie, ainsi que les principales anomalies du sein, en particulier les microcalcifications. Nous abordons également la classification BI-RADS et les notions essentielles d'apprentissage automatique et profond, notamment les réseaux de neurones convolutifs, le transfert d'apprentissage et l'architecture YOLO.

1.2 L'imagerie mammaire :

L'imagerie mammaire est l'utilisation de différentes techniques d'imagerie pour évaluer et diagnostiquer les affections mammaires, en particulier le cancer du sein.

1.2.1 Mammographie :

La mammographie est une radiographie du sein. C'est l'exploration du sein la plus ancienne et la plus souvent réalisée. C'est également la seule technique reconnue pour le dépistage du cancer du sein. Elle utilise de faibles doses de rayons X. c'est pourquoi tout risque de grossesse doit être être signalé au praticien ou radiologue [1].

Une mammographie de dépistage permet de découvrir des cancers du sein à un stade précoce.

Cet examen est recommandé tous les 2 ans chez la femme [1].

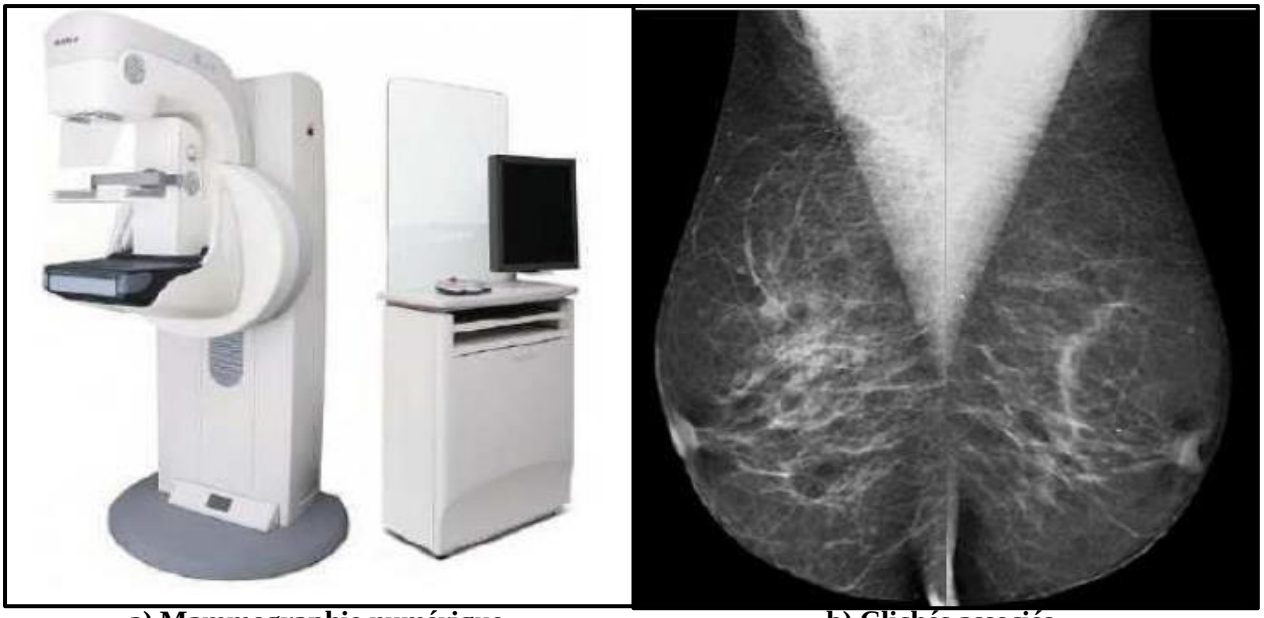
Selon Bermond, près de 40% des cancer « ratés » à la mammographie sont dus à une mauvaise qualité des clichés. Il s'agit le plus souvent d'un défaut de positionnement du sein [2].

1.1.1.1. Du film radiologique aux clichés numérisés :

Il y a deux méthodes de reproduction de l'image mammographique à partir du film radiologique :

a) **La numérasiatation indirecte :** s'effectuer à partir d'un film et une caméra CCD ou d'un numérisateur de film à balayage laser (scanner optique) [3]. Cependant, La numérisation d'une radiographie ne peut jamais révéler plus de détails que l'image originale. Elle conserve en outre toutes les imperfections et les limitations de l'image initiale (bruit, dynamique, artefacts de développement, etc.).

b) **La mammographie numérique :** la technique de la mammographie numérique remplace le système classique d'imagerie par un détecteur qui transforme les rayons X en signaux électroniques. Cette méthode permet une plus grande plage dynamique que le couple écran-film utilisé auparavant. Cette technologie présente de multiples avantages :



a) Mammographie numérique **b) Clichés associés**

Figure 1 : Mammographie numérique (MN) et les clichés associés [3].

Dataset	**Vindr-Mammo**
Nombre d'images	**20000**
Type	**Mammographies FFDM**
Annotation	**Bounding box**
Classes	**Masse bénigne/maligne**
Split	**Train 70% /Val 10% /Test 20%**

Table 1 represente la description du dataset Vindr-Mammo

1.3 Les anomalies mammaires :

Le cancer du sein est une tumeur maligne qui est née par la multiplication anormale et, anarchique des cellules d'un tissu. On cite parmi ces anomalies mammaires:

- Les masses ;
- Les calcifications ;
- Les distorsions architecturales ;
- Les densités asymétriques

1.4 Intelligence artificielle et Machine Learning :

L'intelligence artificielle (IA) regroupe l'ensemble des techniques permettant aux machines de réaliser des tâches nécessitant normalement l'intelligence humaine.

Le Machine Learning est une branche de l'intelligence artificielle permettant aux systèmes d'apprendre automatiquement à partir des données sans être explicitement programmés.

Le Deep Learning représente une sous-catégorie du Machine Learning basée sur les réseaux de neurones artificiels profonds.

Ces techniques sont aujourd'hui largement utilisées dans :

- la vision par ordinateur ;
- la reconnaissance d'images ;
- l'analyse médicale.

1.5 Les calcifications mammaires :

Ce sont des dépôts de sels de calcium qui apparaissent sous forme de points "blancs et brillants" à la mammographie; en fonction de leur taille, on parle de micro ou macrocalcifications. Ces dernières, sont souvent bénignes alors que les microcalcifications demandent plus d'attention. Le diagnostic dépend de leur position dans le sein, de leur arrangement géométrique (linéaire, en forme d'agrégats, ...) et du nombre de microcalcifications dans un amas.

Selon Le Gal [4], on distingue 5 types de microcalcifications (cf. Figure1), les calcifications de type I (rondes et planes), celles de type II (images régulières), celles de type III (images en grains de sel), les calcifications de type IV (points irréguliers) et celles de type V (formes vermiculaires).



(a) Type I (b) Type II (c) Type III (d) Type IV (e) Type

Figure 2: différentes formes des microcalcifications [4]

Parmi les caractéristiques importantes des calcifications (cf. Figure 2), on trouve:

- **La taille des microcalcifications :** les grosses microcalcifications sont le plus souvent bénignes et ont un plus fort contraste que les calcifications malignes ;
- **La forme :** les microcalcifications rondes ou ovales ont plus de chances d'être bénignes que les microcalcifications irrégulières, en "pierres écrasées", ou linéaires ou branchées, qui ont plus de risque d'être malignes.

- **l'orientation :** les amas malins ont tendance à avoir une forme triangulaire ou losangique orientée vers le mamelon ;
- **Le nombre** : plus il y a de microcalcifications dans un amas, plus il est suspect.
- **Le polymorphisme *:*** si les microcalcifications dans un amas ont des formes variées, l'amas a plus de chance d'être malin.



(a) (b)

Figure 3 : (a) Mammographie montrant un foyer de microcalcifications. (b) Agrandissement de la région contenant l'amas de microcalcifications

1.6 Classification BI-RADS :

La classification BI-RADS (Breast Imaging-Reporting And Data System) est utilisée par les radiologistes lors de la mammographie, de l'échographie et de l'IRM pour définir les anomalies vues et permettre de savoir ce qui doit être fait par la suite : retour au dépistage, suivi rapproché ou biopsie [5].

1		Normal	-
2		Anomalie bénigne	-
3		Anomalie probablement bénigne	Suivi dans 6 mois
4	4a	Anomalie Faiblement suspecte	Biopsie
	4b	Anomalie Modérément suspecte	
	4c	Anomalie Fortement suspecte	
5		Anomalie extrêmement suspecte	Biopsie
6		Cancer prouvé par la biopsie	Chirurgie

Table 2 représentant le lexique BIRADS

1.7 Classification automatique des images :

1.7.1 Définition :

La classification automatique des images consiste à attribuer automatiquement une classe à une image à l‟aide d‟un système de classification sans intervention humaine. On retrouve ainsi la classification d‟objets, de scènes, de textures, la reconnaissance de visages, d‟empreintes digitale et de caractères [6].

1.7.2 Objectif :

L‟objectif de la classification d'images est d'élaborer un système capable d'affecter une classe automatiquement à une image comme par exemple (classe animaux, classe humaine, classe transport, etc.). Ainsi, ce système permet d'effectuer une tâche d'expertise qui peut s'avérer coûteuse à acquérir pour un être humain en raison notamment de contraintes physiques comme la concentration, la fatigue ou le temps nécessaire à un volume important de données images.

1.8 L'apprentissage automatique :

Le vaste domaine d'apprentissage automatique (en anglais, machine Learning) vise notamment à reconnaitre des objets et à faire des prévisions à partir de données existantes. Il sert aussi à gérer des tâches de classification et de segmentation qui servent à détecter des corrélations entre paramètres et événements. Machine Learning a pour objectif de construire des modèles mathématiques permettant de prédire une sortie étant donné un signal d'entrée. Les réseaux de neurones sont des outils permettant d'apprendre ces modèles. Nous pouvons distinguer différentes familles d'apprentissage (supervisé, semi-supervisé, non-supervisé, par renforcement...) [7].

1.9 L'apprentissage profond "Deep Learning" :

L'apprentissage profond « Deep Learning » est une sous-branche du Machine Learning, il se caractérise par l'utilisation d'algorithmes basés sur une architecture de réseaux de neurones et nécessite pour une exploitation totale de ses capacités de larges quantités de données.

L'utilisation de l'apprentissage profond s'est répandue à divers domaines d'application, surpassant les modèles précédents basés sur des algorithmes de « Machine Learning » classiques (KNN, Random Forest, SVM etc..).

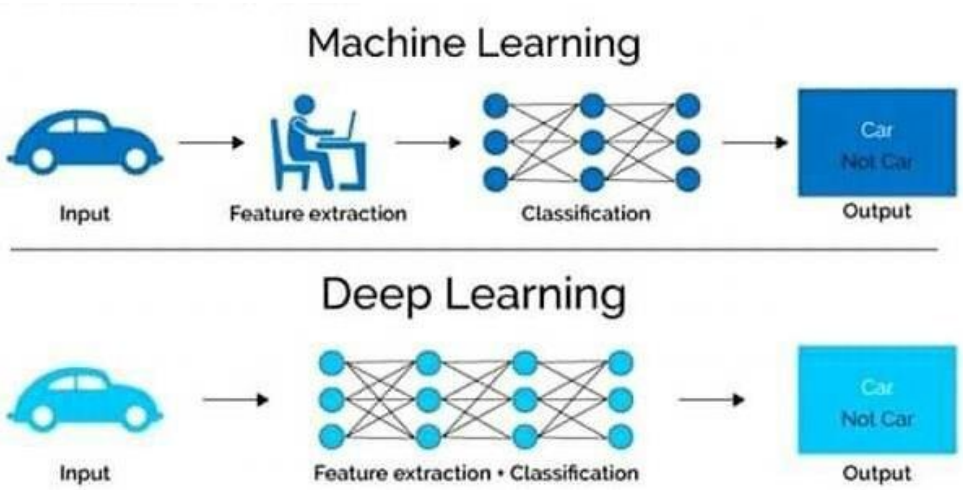


Figure 4 : La différence entre l'apprentissage profond et l'apprentissage automatique

1.10 Les réseaux de neurones convolutifs :

L'idée de base d'un réseau de neurones convolutifs (CNN) vient de l'inspiration de la façon dont notre cerveau fonctionne avec les neurones connectés de la même manière, une CNN est construite. Les neurones sont organisés en couches connectées à d'autres couches de neurones où l'un peut activer l'autre. La façon dont ces neurones sont connectés dépend de la structure et la fonction spécifique de chaque couche [8]. Cela crée des modèles de programmation puissants qui permettant notamment la reconnaissance d'images, en attribuant automatiquement à chaque image fournie en entrée une étiquette correspondant à sa classe d'appartenance, après l'extraction des attributs pour chaque région de cette image (*figure 3*).

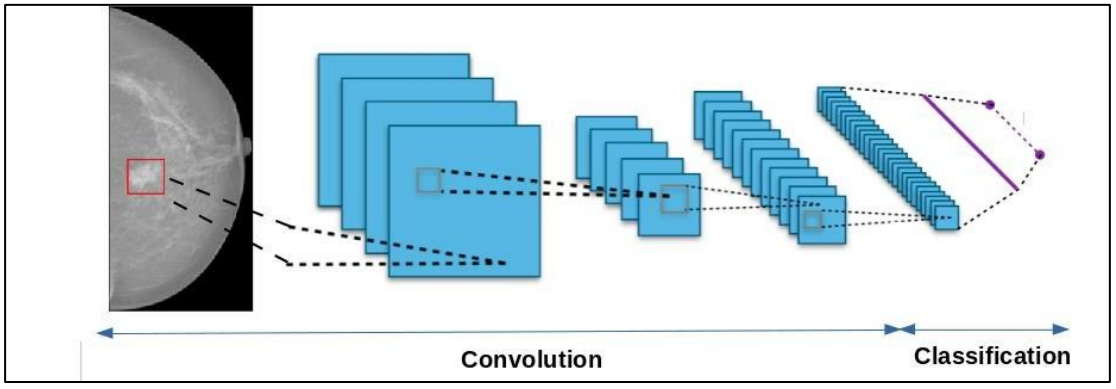


Figure 5 : L'Architecture d'un réseau de neurones convolutifs.

1.10.1 La convolution :

La convolution est un outil mathématique simple qui est très largement utilisé pour le traitement d'image. Sur une image, son effet s'assimile à un filtrage. On définit une taille de fenêtre qui va se balayer à travers toute l'image. Au début de la convolution, la fenêtre sera positionnée tout en haut à gauche de l'image puis elle va se déplace progressivement de la gauche vers la droite d'un certain nombre de cases défini au préalable (le pas) jusqu'à arriver au bout de l'image (*Figure 4).*

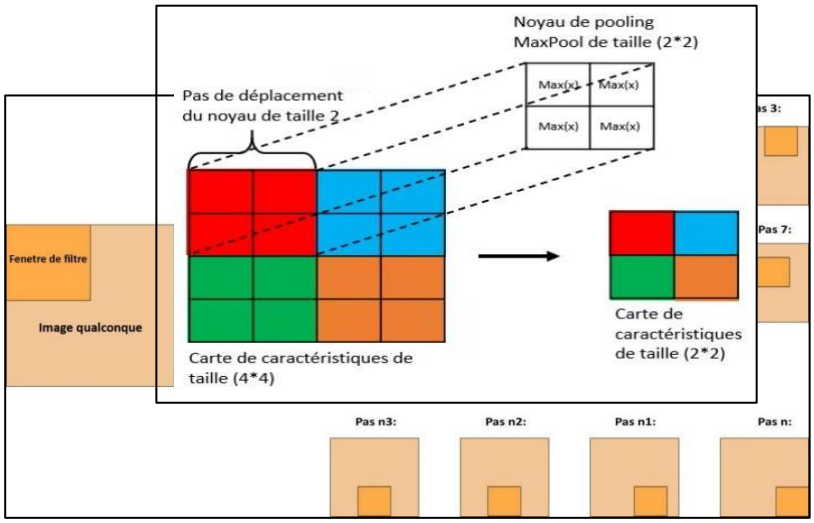


Figure 6 : La forme générale de l'Architecture d'un réseau de neurones convolutifs.

À chaque portion d'image rencontrée, un calcul de convolution s'effectue permettant d'obtenir en sortie une carte d'activation ou feature map qui indique la localisation des paramètres dans l'image : plus la feature map est élevée, plus la portion de l'image balayée ressemble aux paramètres *(figure 2.13).*

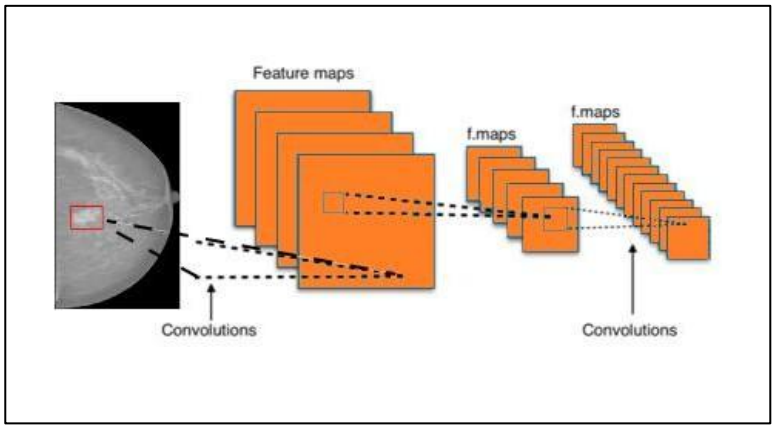


Figure 7 : La réduction de la taille de la carte de caractéristiques (feature map).

1.10.2 Le Pooling :

Le Pooling est un processus de discrétisation basé sur des échantillons. Son objectif est de sous-échantillonner une représentation d'entrée (image, matrice de sortie de couche cachée, etc.) en réduisant sa dimension. Il est utilisé pour réduire la taille ou la dimensionnalité de feature map afin de réduire le nombre de paramètres nécessaires à l'entraînement.
Il existe trois types de Pooling que nous pouvons appliquer : Max, Moyenne et Somme. Le Max-Pooling qui fournit une invariance par petites translations : si une petite translation ne modifie pas le maximum de la région balayée, le maximum de chaque région restera le même et donc la nouvelle matrice créée restera identique *(Figure 2.14).*

Figure 8 : Une opération de pooling avec un noyau MaxPool de taille 2*2 et d'un pas de 2 [9]

1.10.3 Les couches entièrement connectées :

Ces couches sont placées en fin d'architecture de CNN. Cette partie constituée de couches entièrement connectées à tous les neurones de sorties appelées perceptron multicouches (fully

connected layers). Le rôle de cette partie est de combiner les caractéristiques après avoir reçu un vecteur en entrée, la couche FC applique successivement une combinaison linéaire puis une fonction d'activation afin de classer l'entrée « image », à la fin elle renvoie en sortie un vecteur de taille correspondant au nombre de classes dans lequel chaque composant représente la probabilité d'appartenance d'une entrée « image » à une classe *(figure 2.15).[10]*

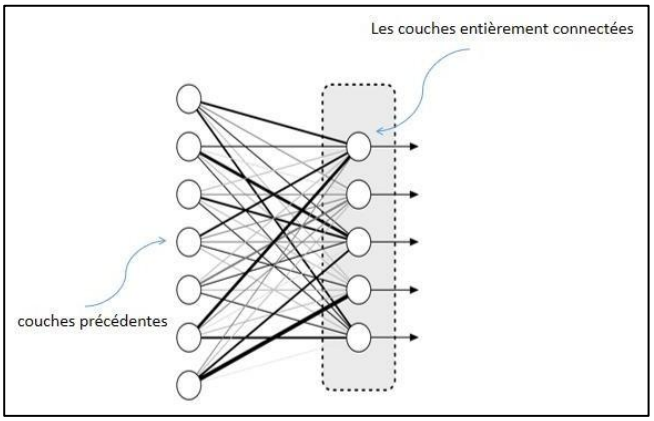


Figure 9 : La couche entièrement connectée [11]

1.11 Transfert des connaissances « Tranfert learning » :

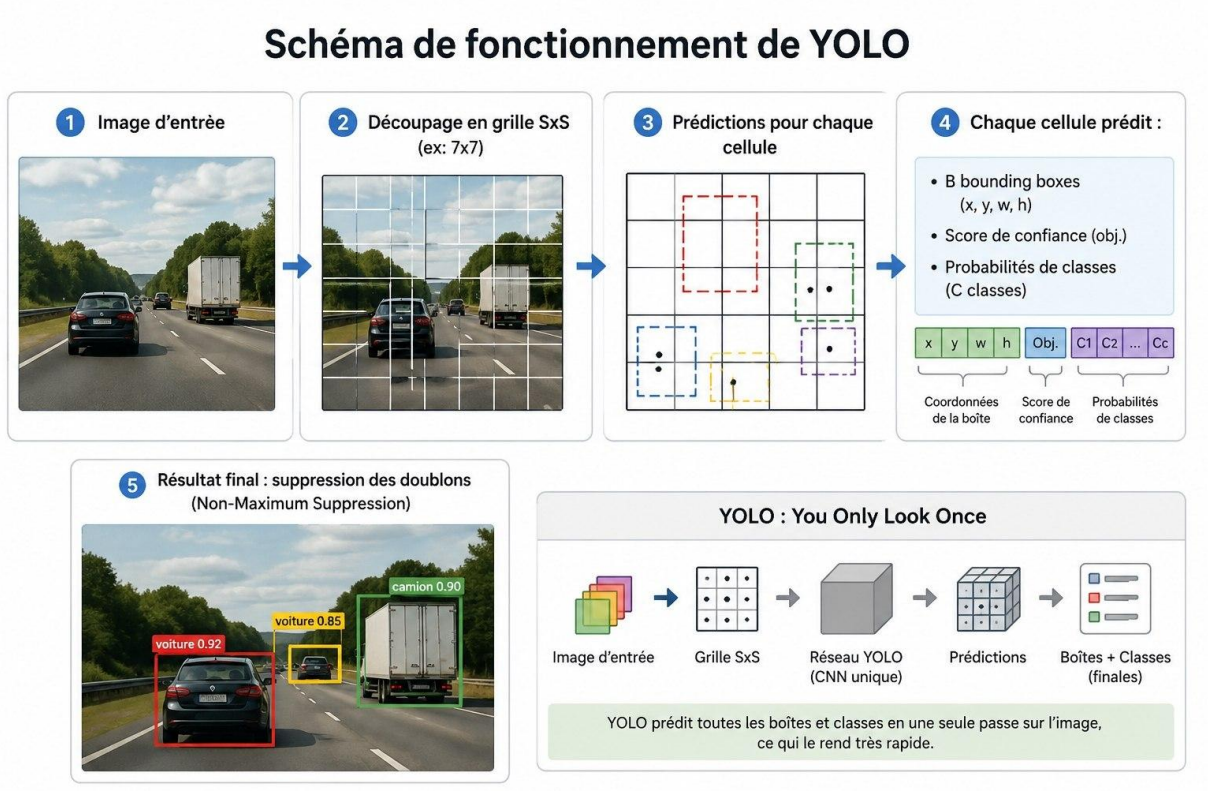
Un réseau neuronal profond comportant plusieurs couches cachées est coûteux à entraîner. Les modèles complexes nécessitent plusieurs machines équipées de GPU onéreux et leur entraînement peut prendre des semaines. L'apprentissage par transfert est une approche courante en apprentissage profond, consistant à utiliser des modèles pré-entraînés comme point de départ pour le nouveau réseau neuronal. Ses principaux domaines d'application sonla vision par ordinateur et le traitement automatique du langage naturel. Différentes variantes de réseaux pré-entraînés sont disponibles, chacune présentant une architecture, une vitesse, une taille, des avantages et des inconvénients qui lui sont propres. Parmi les modèles les plus courants, on trouve VGGNet et ResNet.
Dans cette étude, le modèle InceptionResNetV2, pré-entraîné sur le jeu de données ImageNet, a été utilisé. Les mammographies, en tant qu'images, diffèrent des images naturelles par le nombre de canaux et la complexité intrinsèque de leur structure. Par conséquent, le modèle InceptionResNetV2 pré-entraîné a été utilisé uniquement pour obtenir les cartes de caractéristiques nécessaires à la classification. L'entraînement du système de classification des microcalcifications en trois catégories a été réalisé sur le jeu de données CBIS-DDSM.

1.12 L'architecture YOLO :

De nombreux modèles de détection d'objets prennent et traitent l'image plusieurs fois pour pouvoir détecter tous les objets présents dans les images, pour cela l'idée de l'algorithme YOLO est née. En 2015 Joseph Redmons a développé l'architecture de cet algorithme de façon à ce qu'il n'applique qu'une seule phase vers l'avant à l'ensemble de l'image et prédit les rectangles englobantes et leurs probabilités de classe, d'où la nomination « You Only Look Once » (vous ne regardez qu'une seule fois). Cela rend cet algorithme de détection d'objets en temps réel ultra-rapide. L'implémentation principale de Redmons YOLO est basée sur Darknet, qui est un cadre de recherche très flexible écrit dans des langages de bas niveau et a produit une série des meilleurs détecteurs d'objets en temps réel en vision par ordinateur : YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOV5, YOLOv6, YOLOv7 [12].
Le fonctionnement de l'algorithme YOLO peut être expliqué à travers trois techniques :

- **Blocs résiduels :** Tout d'abord, l'image est divisée en différentes grilles. Chaque grille est de dimension S x S. La procédure de conversion d'une image d'entrée en grilles est représentée dans l'image suivante. Chaque cellule de la grille détectera les objets qui y apparaissent.
- **Régression de boîte englobante 'Bounding box regression' :** le cadre englobant est un contour qui met en évidence un objet dans une image avec certains attributs tels que la largeur (bw), la hauteur (bh), la classe (par exemple, personne, voiture, etc.), ceci est représenté par la lettre c. Centre de la boîte englobante (bx,by). YOLO utilise une seule régression de boîte englobante pour prédire la hauteur, la largeur, le centre et la classe des objets.

- **Intersection sur Union (IOU) :** L'intersection sur l'union (IOU) est un mécanisme de détection d'objet qui décrit comment les boîtes se chevauchent 'boxes overlap'. YOLO utilise IOU pour entourer parfaitement les boîtes de sortie parfaites pour les objets. Chaque cellule de la grille est chargée de prédire les boîtes englobantes et leurs scores de confiance. Si la boîte englobante prédite est la même que la boîte réelle, alors l'IOU est égal à 1. Cette technique peut éliminer les boîtes englobantes qui ne sont pas égales à la boîte réelle.



1.1.1.2. Figure10 :schema de yolo

1.1.1.3.

1.13 Métriques d'évaluation :

L'évaluation des performances des modèles de Deep Learning constitue une étape essentielle dans les applications médicales. Les métriques d'évaluation permettent de mesurer la qualité

des prédictions réalisées par le modèle en comparant les résultats obtenus avec les annotations réelles.

Dans notre travail, nous avons principalement utilisé les métriques suivantes :

- ***Precision ;***
- ***Recall (Sensibilité) ;***
- ***Intersection over Union (IoU) ;***
- ***Dice Coefficient ;***
- ***mAP50 ;***
- ***mAP50-95.***

Ces métriques permettent d'évaluer la capacité du modèle YOLOv11 à détecter et localiser correctement les masses mammaires dans les images mammographiques.

Conclusion :

Dans ce chapitre, nous avons présenté les notions fondamentales liées au cancer du sein, à la mammographie et aux anomalies mammaires. Nous avons également abordé les concepts de base du Deep Learning, des réseaux de neurones convolutifs ainsi que le modèle YOLO utilisé dans notre étude.

Enfin, nous avons présenté les principales métriques d'évaluation utilisées dans les systèmes de détection d'objets médicaux telles que l'IoU, le Dice Coefficient et le mAP.

Ces éléments constituent la base théorique nécessaire pour la suite de ce travail consacré à la détection et à la localisation des masses mammaires.

CHAPITRE 02

Approche Méthodologique

2.1 Introduction :

Dans ce chapitre, nous présentons la méthodologie adoptée pour la détection et la localisation des masses mammaires dans les images mammographiques à l'aide des techniques d'apprentissage profond. Notre approche repose principalement sur l'utilisation du modèle YOLOv11 appliqué au dataset médical VinDr-Mammo.

L'objectif principal de ce travail est d'évaluer les performances du modèle YOLOv11 dans deux scénarios différents : un entraînement sans Transfer Learning et un entraînement avec Transfer Learning utilisant des poids pré-entraînés sur le dataset COCO.

Dans ce chapitre, nous décrivons les datasets utilisés, les outils logiciels employés, les techniques de prétraitement des données, l'architecture YOLOv11 ainsi que les métriques d'évaluation utilisées pour mesurer les performances du modèle.

2.2 Datasets utilisés :

2.2.1 Dataset CBIS-DDSM :

Le dataset **CBIS-DDSM** (Curated Breast Imaging Subset of DDSM) est une version améliorée et réorganisée du célèbre dataset DDSM destiné à l'analyse des mammographies.

Ce dataset contient des mammographies annotées comprenant plusieurs anomalies mammaires telles que :

- les masses ;
- les microcalcifications ;
- les asymétries ;
- les distorsions architecturales.

Les images sont annotées par des radiologues experts avec des régions d'intérêt (ROI) permettant de localiser précisément les anomalies.

Le dataset **CBIS-DDSM** est largement utilisé dans les travaux de recherche en intelligence artificielle appliquée au diagnostic du cancer du sein grâce à :

- la qualité de ses annotations ;
- sa disponibilité publique ;
- la diversité des anomalies ;
- son utilisation fréquente dans la littérature scientifique.

Dans notre étude, ce dataset a été étudié dans le cadre de l'analyse bibliographique et de la comparaison avec les travaux existants.

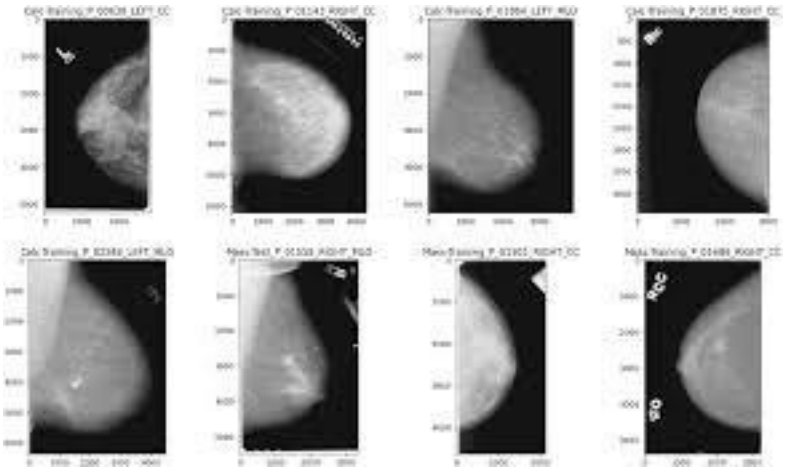


Figure 11 : Exemple de datsset CBIS-DDSM

2.2.2 Dataset VinDr-Mammo :

Le dataset VinDr-Mammo est une base de données récente spécialisée dans les mammographies numériques FFDM (Full-Field Digital Mammography). Il a été développé afin de fournir des données médicales de haute qualité pour les applications d'intelligence artificielle.

Le dataset contient plusieurs milliers d'images mammographiques annotées par des radiologues expérimentés. Les annotations comprennent différentes anomalies mammaires telles que :

- les masses ;
- les calcifications ;
- les asymétries ;
- les distorsions architecturales.

Dans notre travail, nous avons utilisé principalement les annotations correspondant aux masses mammaires afin d'entraîner le modèle YOLOv11 pour la tâche de détection.

Le dataset a été divisé en :

- ensemble d'entraînement ;
- ensemble de validation ;
- ensemble de test.

Les annotations ont été converties au format YOLO afin d'être compatibles avec le framework Ultralytics.

Le choix du dataset VinDr-Mammo est justifié par :

- la qualité des annotations ;
- la présence des bounding boxes ;
- la qualité des mammographies numériques ;

- son adaptation aux tâches de détection d'objets médicaux.

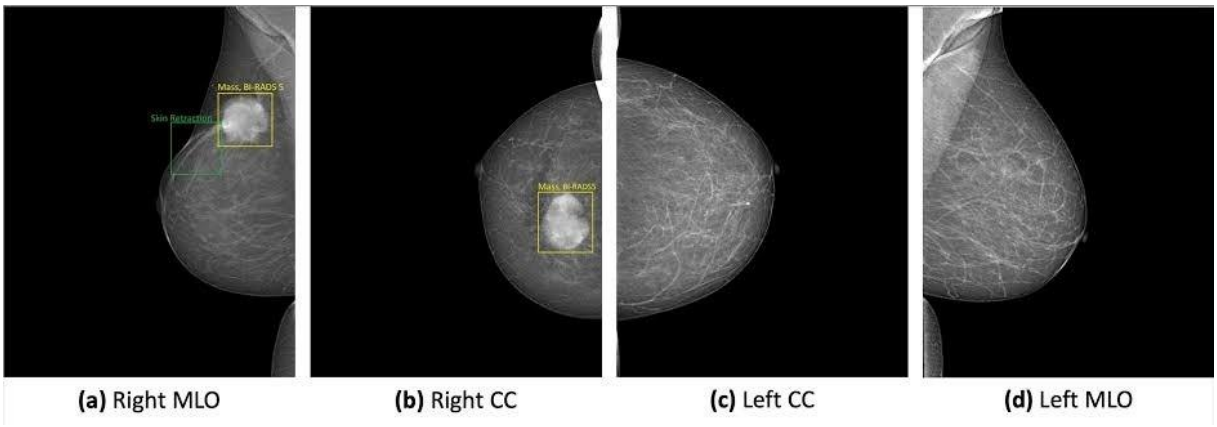


Figure12: Exemple de datasetVinDR-Mammo

2.3 Prétraitement des données :

Le prétraitement des données est une étape essentielle dans les systèmes de Deep Learning. Avant l'entraînement du modèle, plusieurs opérations ont été appliquées afin d'améliorer la qualité des données et d'optimiser les performances du modèle.

Les principales opérations réalisées sont :

- redimensionnement des images ;
- normalisation des données ;
- conversion des annotations vers le format YOLO ;
- suppression des données incorrectes ;
- vérification des annotations ;
- augmentation des données (Data Augmentation).

Les techniques de Data Augmentation utilisées sont :

- rotation ;
- flip horizontal ;
- variation de luminosité ;
- zoom.

Ces opérations permettent :

- d'augmenter la diversité des données ;
- d'améliorer la généralisation du modèle ;
- de réduire le phénomène de surapprentissage (Overfitting).

2.4 Environnement de travail :

2.4.1 Kaggle :

Kaggle est une plateforme en ligne spécialisée dans la science des données, le Machine Learning et l'intelligence artificielle. Elle permet aux chercheurs, étudiants et développeurs de créer, exécuter et partager des projets de Deep Learning directement dans un environnement cloud sans nécessiter de configuration locale.

Kaggle offre plusieurs services importants tels que :

- des datasets publics ;
- des notebooks Python interactifs ;
- des ressources GPU gratuites ;
- des compétitions en intelligence artificielle ;
- une communauté de chercheurs et développeurs.

Dans notre travail***, Kaggle*** a été utilisé comme environnement principal pour l'entraînement et l'évaluation du modèle YOLOv11 grâce à ses ressources GPU et sa compatibilité avec les frameworks de Deep Learning comme PyTorch.

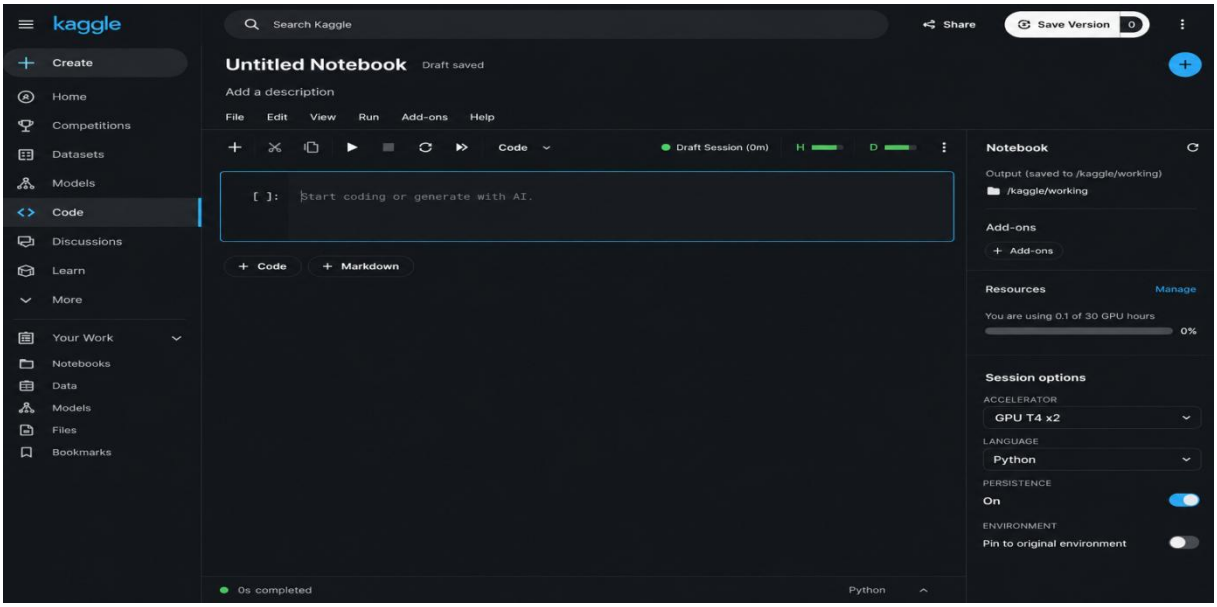


Figure 13: Interface de la plateforme Kaggle utilisée pour l'entraînement des modèles de Deep Learning .

2.4.2 PyTorch :

PyTorch est un framework open-source de Deep Learning développé par Meta AI. Il est largement utilisé dans les applications de vision par ordinateur grâce à sa simplicité et sa flexibilité.Dans notre projet, PyTorch a été utilisé pour l'implémentation et l'entraînement des modèles.

2.4.3 Ultralytics YOLO :

Ultralytics est une bibliothèque moderne permettant l'utilisation des modèles YOLO pour les tâches de détection d'objets.

Cette bibliothèque facilite :

- l’entraînement ;
- l’évaluation ;
- l’inférence ;
- la visualisation des résultats.

Dans notre étude, nous avons utilisé Ultralytics pour entraîner le modèle YOLOv11 sur le dataset VinDr-Mammo.

2.5 Architecture YOLOv11 :

YOLOv11 (You Only Look Once version 11) est un modèle avancé de détection d’objets basé sur les réseaux de neurones convolutifs (***CNN***). Il appartient à la famille des modèles **YOLO** reconnus pour leur rapidité et leurs hautes performances dans les applications de vision par ordinateur.

Contrairement aux méthodes classiques qui analysent l’image en plusieurs étapes, **YOLO** traite l’image une seule fois afin de détecter directement les objets présents ainsi que leurs coordonnées.

Le principe de fonctionnement de **YOLO** repose sur la division de l’image en plusieurs cellules de grille. Chaque cellule est responsable de la prédiction :

- des coordonnées des objets ;
- des classes ;
- des scores de confiance.

YOLOv11 est composé principalement de trois parties :

- **Backbone :**

Le Backbone est chargé de l’extraction des caractéristiques importantes de l’image. Il permet d’identifier les contours, textures et formes importantes nécessaires à la détection des anomalies mammaires.

- **Neck :**

Le Neck assure la fusion des informations provenant de différents niveaux de caractéristiques. Cette étape améliore la capacité du modèle à détecter les objets de différentes tailles.

- **Head :**

Le Head est responsable de la prédiction finale des bounding boxes, des classes et des scores de confiance.

Le modèle ***YOLOv11*** présente plusieurs avantages :

- rapidité d’exécution ;
- haute précision ;
- bonne généralisation ;
- détection en temps réel ;

- capacité à traiter des images complexes.

Dans notre travail, ***YOLOv11*** a été utilisé pour détecter automatiquement les masses mammaires présentes dans les mammographies numériques.

2.6 Transfer Learning :

Le Transfer Learning est une technique de Deep Learning consistant à réutiliser un modèle pré-entraîné sur un grand dataset afin de résoudre une nouvelle tâche similaire.

Cette approche permet :

- de réduire le temps d'entraînement ;
- d'améliorer les performances ;
- de limiter le besoin de grandes quantités de données.

Dans le domaine médical, les datasets annotés sont souvent limités. Le Transfer Learning permet donc d'exploiter les connaissances déjà apprises par le modèle sur des datasets plus larges.

Dans notre étude, nous avons utilisé des poids pré-entraînés sur le dataset COCO avant d'effectuer un Fine-Tuning sur le dataset VinDr-Mammo.

Deux approches ont été étudiées :

- entraînement sans Transfer Learning ;
- entraînement avec Transfer Learning.

L'objectif est d'évaluer l'impact réel du Transfer Learning sur les performances du modèle YOLOv11 dans la détection des masses mammaires.

2.7 Dataset COCO :

Le dataset ***COCO*** (Common Objects in Context) est l'un des benchmarks les plus populaires dans le domaine de la vision par ordinateur,il contient :

- plus de 330 000 images ;
- environ 1,5 million d'objets annotés ;
- 80 catégories d'objets.

Le dataset ***COCO*** est largement utilisé pour le pré-entraînement des modèles de détection d'objets grâce à :

- sa grande diversité visuelle ;
- la qualité de ses annotations ;
- la variété des objets présents.

Dans notre travail, les poids pré-entraînés sur ***COCO*** ont été utilisés comme point de départ pour l'entraînement du modèle ***YOLOv11***.

Cette approche permet au modèle d'apprendre des caractéristiques générales avant d'être spécialisé dans les mammographies médicales.



Figure 14 : Exemple de dataset COCO

2.8 Paramètres d'entraînement :

Les principaux hyperparamètres utilisés lors de l'entraînement sont présentés dans le tableau suivant :

Hyperparametre	valeur
Modèle	**YOLOv11**
Nombre d'époques	**500**
Batch size	**8**
Taille des images	**1280**
Optimiseur	**AdamW**
Patience	**50**
Device	**GPU (0,1)**
Framework	**pyTorch**

Tableau 3 : hyperparametres lors de l'entrainement

2.9 Métriques d'évaluation :

L'évaluation des performances du modèle est une étape essentielle dans les systèmes de détection médicale. Plusieurs métriques ont été utilisées afin de mesurer la qualité des prédictions du modèle YOLOv11.

2.9.1 Intersection over Union (IoU) :

L'Intersection over Union (IoU) est une métrique utilisée pour mesurer le degré de chevauchement entre la boîte prédite par le modèle et la boîte réelle annotée dans l'image.

Elle est calculée selon la formule suivante :

IoU = Aire de l'intersection / Aire de l'union

Une valeur élevée de ***IoU*** indique une bonne localisation de l'objet détecté.

Lorsque l'***IoU*** est égal à 1, cela signifie que la boîte prédite correspond parfaitement à la vérité terrain.

L'***IoU*** joue un rôle fondamental dans le calcul des métriques mAP50 et mAP50-95.

2.9.2 Dice Coefficient :

Le Dice Coefficient est une métrique utilisée principalement dans les applications médicales afin de mesurer la similarité entre deux régions.

Il est défini par :

Dice = (2 × Intersection) / (A + B)

où :

- A représente la région réelle .
- B représente la région prédite.

Une valeur proche de 1 indique une forte similarité entre les deux régions.

Le ***Dice*** Score est particulièrement important dans l'évaluation des performances des modèles médicaux.

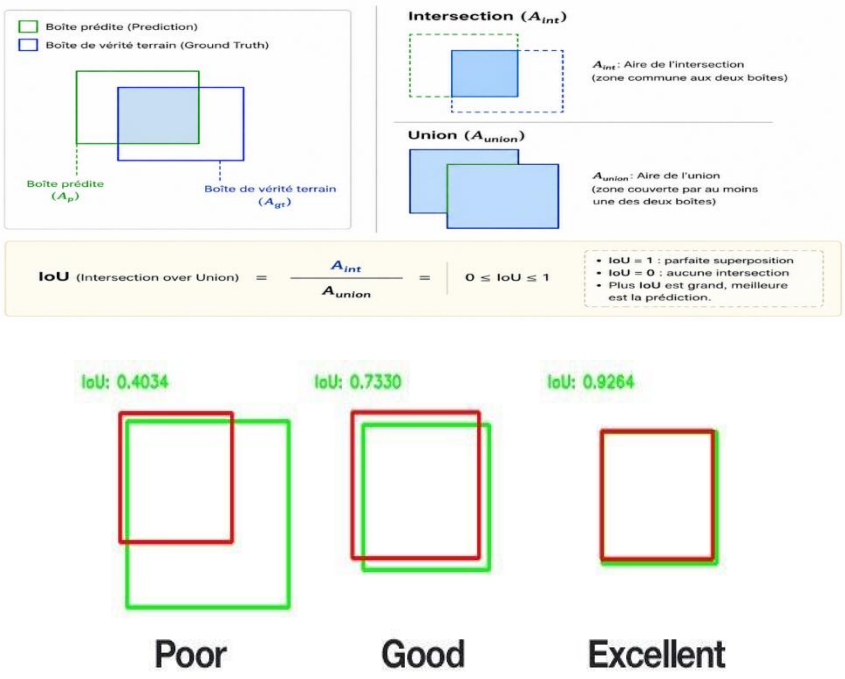


Figure 15 : IoU evaluation examples

2.9.3 mAP50 :

Le mAP50 (mean Average Precision at IoU = 0.50) représente la précision moyenne calculée lorsque le seuil IoU est fixé à 0.50.

Une détection est considérée correcte si le chevauchement entre la boîte prédite et la boîte réelle dépasse 50 %.

Cette métrique permet d'évaluer les performances générales du modèle.

2.9.4 mAP50-95 :

Le mAP50-95 correspond à la moyenne des précisions calculées pour plusieurs seuils IoU allant de 0.50 jusqu'à 0.95 avec un pas de 0.05.

Cette métrique est plus stricte que le mAP50 car elle exige une meilleure précision de localisation.

Le mAP50-95 fournit une évaluation plus complète et plus robuste des performances du modèle.

2.10 Les approches proposées :

2.10.1 Approche sans Transfer Learning :

Dans cette première approche, le modèle YOLOv11 a été entraîné à partir de zéro sans utiliser de poids pré-entraînés.

Les poids du réseau ont été initialisés aléatoirement et le modèle a appris uniquement à partir des images du ***dataset VinDr-Mammo***.

Cette approche permet d'évaluer la capacité intrinsèque du modèle à apprendre directement les caractéristiques des masses mammaires.

2.10.2 Approche avec Transfer Learning :

Dans cette seconde approche, nous avons utilisé des poids pré-entraînés sur le dataset COCO avant d'effectuer un Fine-Tuning sur le ***dataset VinDr-Mammo***.

Cette stratégie permet :

- d'améliorer les performances ;
- d'accélérer l'apprentissage ;
- de réduire le temps d'entraînement.

L'objectif est de comparer les performances des deux approches afin d'évaluer l'apport réel du Transfer Learning.

Conclusion :

Dans ce chapitre, nous avons présenté la méthodologie adoptée pour la détection des masses mammaires à l'aide du modèle YOLOv11. Nous avons décrit les datasets utilisés, les outils logiciels employés, les techniques de prétraitement des données ainsi que les différentes approches d'entraînement.

Nous avons également présenté les métriques d'évaluation utilisées pour mesurer les performances du modèle.

Le chapitre suivant sera consacré à la présentation et à l'analyse des résultats expérimentaux obtenus.

Chapitre 03 :

Résultats et discussions

Dans ce chapitre, nous présentons une analyse comparative des résultats obtenus à partir des différentes expériences réalisées avec les modèles YOLO appliqués à la détection des anomalies mammaires sur le dataset VinDr-Mammo. L'objectif principal de cette étude est d'évaluer l'impact du Transfer Learning sur les performances des modèles de détection, en comparant un entraînement classique (from scratch) avec un entraînement basé sur des poids pré-entraînés sur le dataset COCO.
Les performances ont été évaluées à l'aide des métriques standards utilisées en détection d'objets, notamment le mAP50 et le mAP50-95, qui permettent de mesurer la qualité de localisation et de classification des objets détectés.

3.2 Comparaison des modèles utilisés :

Dans cette étude, deux variantes de l'architecture YOLO ont été exploitées :
YOLO11n avec Transfer Learning
YOLO26n avec Transfer Learning
Les deux modèles ont été entraînés sur le dataset VinDr-Mammo après initialisation avec des poids pré-entraînés sur le dataset COCO. Cette stratégie permet au modèle de bénéficier de connaissances visuelles générales apprises sur un très grand nombre d'images avant d'être spécialisé dans le domaine médical.
Le tableau suivant présente les résultats obtenus :

Table 4 represent la differnce entre yolo11n et yolo26n

Model	mAP50	mAP50-95
Yolo11n avec transfert learning	0.654	0.369
Yolo26n avec transfert learning	0647	0.380

3.3 Analyse des résultats :

Les résultats expérimentaux montrent que les deux modèles obtiennent des performances relativement proches. Le modèle YOLO11n atteint une valeur de 0.654 pour le mAP50, ce qui indique une bonne capacité à détecter correctement les anomalies mammaires présentes dans les images mammographiques. En revanche, le modèle YOLO26n obtient un mAP50 légèrement inférieur avec une valeur de 0.647.
Cependant, en observant la métrique mAP50-95, qui représente une évaluation plus stricte prenant en compte plusieurs seuils d'Intersection over Union (IoU), le modèle YOLO26n affiche de meilleures performances avec une valeur de 0.380, contre 0.369 pour YOLO11n. Cela signifie que YOLO26n produit des boîtes englobantes plus précises et une meilleure localisation des régions suspectes.
Cette différence peut s'expliquer par la complexité plus importante de YOLO26n, lui permettant d'extraire davantage de caractéristiques pertinentes à partir des mammographies. Toutefois, cette amélioration reste modérée, ce qui montre que les deux modèles sont globalement efficaces pour cette tâche de détection médicale.

3.4 Impact du Transfer Learning :

L'utilisation du Transfer Learning a permis d'améliorer les performances des modèles tout en réduisant le temps d'apprentissage. En effet, les poids pré-entraînés sur COCO offrent une base de connaissances visuelles utile pour l'extraction des caractéristiques générales telles que les contours, les textures et les formes.

Dans le domaine médical, les datasets annotés sont souvent limités et difficiles à obtenir. Le recours au Transfer Learning constitue donc une solution efficace pour améliorer la capacité de généralisation des modèles et éviter le surapprentissage (overfitting).

Les résultats obtenus confirment que les modèles pré-entraînés peuvent être adaptés avec succès à la détection des anomalies mammaires malgré la différence importante entre les images naturelles du dataset COCO et les images médicales mammographiques.

3.5 Limites de l'étude :

Malgré les résultats encourageants obtenus dans cette étude, certaines limites doivent être prises en considération :

La taille du dataset VinDr-Mammo reste relativement limitée par rapport aux grands datasets utilisés dans la vision par ordinateur.

Les anomalies mammaires présentent parfois des formes très complexes et difficiles à détecter.

Certaines images contiennent du bruit ou des contrastes faibles pouvant affecter les performances du modèle.

Les ressources matérielles disponibles limitent l'utilisation de modèles plus lourds et plus profonds.

Une augmentation du volume de données ainsi que l'utilisation de techniques avancées d'augmentation de données pourraient améliorer davantage les résultats.

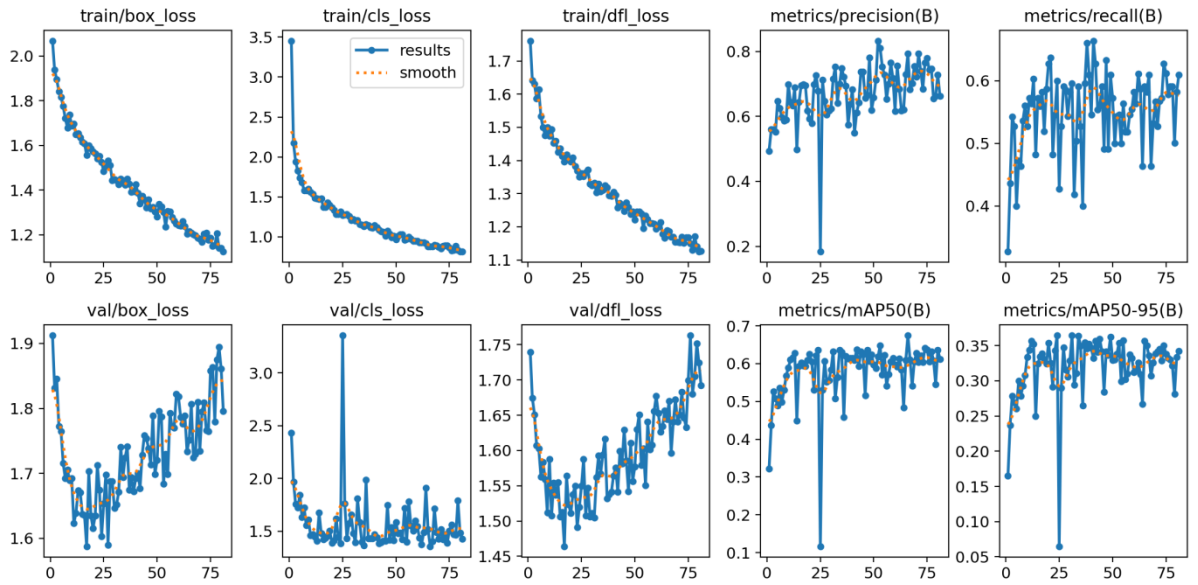


Figure 16 :yolo training metrices

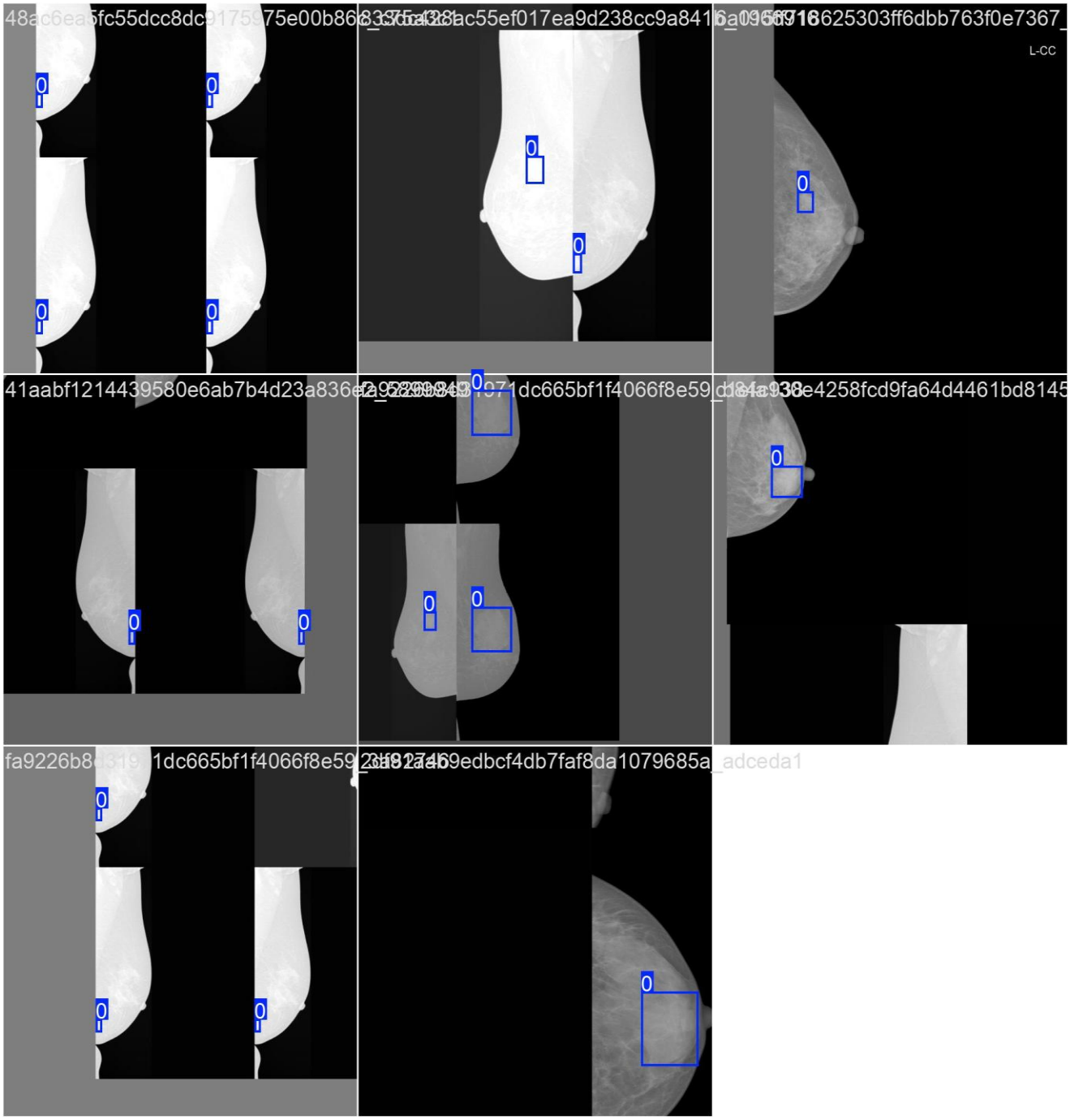
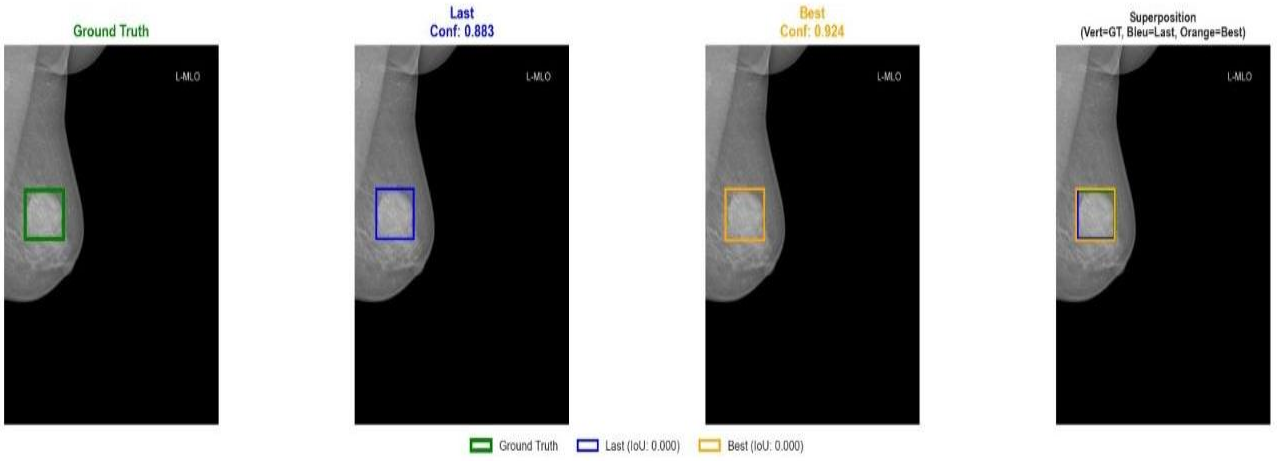
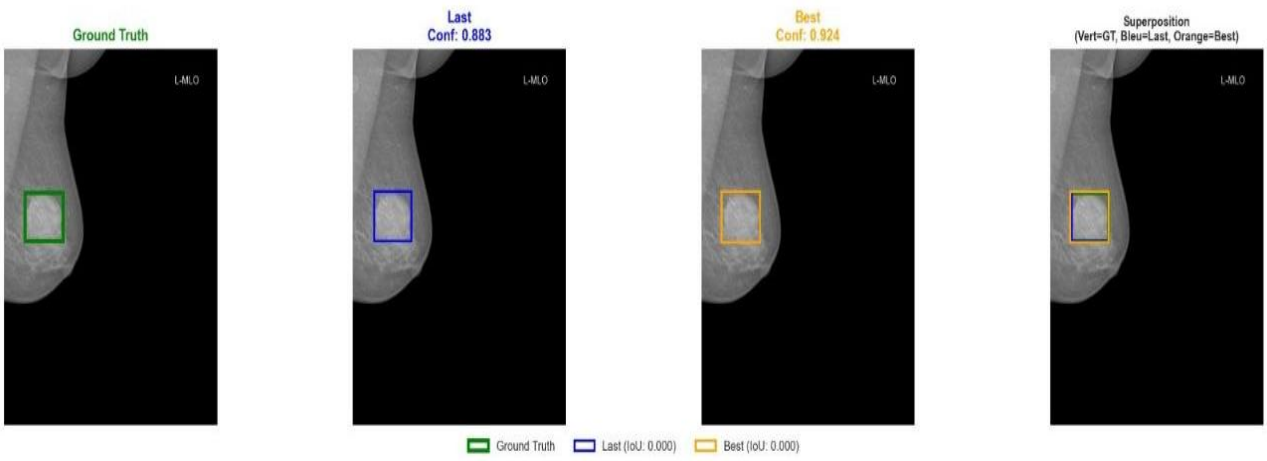


Figure 17 :YOLO detection examples


Ground Truth
Last
Conf: 0.573
Best
Conf: 0.693
Superposition
(Vert=GT, Bleu=Last, Orange=Best)
Ground Truth
Last (IoU: 0.000)
Best (IoU: 0.000)


Ground Truth
Last
Non détecté
Best
Conf: 0.805
Superposition
(Vert=GT, Bleu=Last, Orange=Best)
L-MLO
Ground Truth
Last (IoU: 0.000)
Best (IoU: 0.000)


Ground Truth
Last
Conf: 0.883
Best
Conf: 0.924
Superposition
(Vert=GT, Bleu=Last, Orange=Best)
L-MLO
Ground Truth
Last (IoU: 0.000)
Best (IoU: 0.000)


Ground Truth
Last
Conf: 0.883
Best
Conf: 0.924
Superposition
(Vert=GT, Bleu=Last, Orange=Best)
L-MLO
Ground Truth
Last (IoU: 0.000)
Best (IoU: 0.000)

Conclusion :

Dans ce chapitre, nous avons effectué une analyse comparative des modèles YOLO11n et YOLO26n entraînés avec la technique du Transfer Learning. Les résultats montrent que les deux modèles offrent des performances satisfaisantes pour la détection des anomalies mammaires sur le dataset VinDr-Mammo.

Le modèle YOLO11n présente un meilleur score mAP50, tandis que YOLO26n obtient une meilleure précision globale selon la métrique mAP50-95. Cette étude confirme également l'importance du Transfer Learning dans les applications médicales basées sur l'apprentissage profond, notamment lorsque les données annotées sont limitées.

Enfin, ces travaux ouvrent la voie à de futures améliorations basées sur des architectures plus avancées et des datasets médicaux plus riches afin d'obtenir des systèmes de détection encore plus précis et robustes.

Conclusion Générale :

Le cancer du sein demeure l'une des principales causes de mortalité chez les femmes dans le monde. Son dépistage précoce joue un rôle essentiel dans l'amélioration des chances de guérison et dans la réduction du taux de mortalité. Dans ce contexte, les techniques d'intelligence artificielle et d'apprentissage profond offrent aujourd'hui des solutions prometteuses pour assister les radiologues dans l'analyse des images mammographiques.
Dans ce mémoire, nous nous sommes intéressés à la détection et à la classification des anomalies mammaires à l'aide de modèles d'apprentissage profond. Après avoir présenté les fondements médicaux liés à la mammographie et aux microcalcifications, ainsi que les concepts théoriques de l'apprentissage automatique et profond, nous avons mis en œuvre une approche expérimentale basée sur l'architecture YOLOv11 et la technique du transfert de connaissances.
Les expériences réalisées sur le jeu de données VinDr-Mammo ont permis d'évaluer les performances du modèle dans différentes configurations d'apprentissage. Les résultats obtenus montrent que les méthodes basées sur les réseaux de neurones profonds constituent des outils efficaces pour l'analyse automatique des mammographies et peuvent contribuer à améliorer le processus de dépistage du cancer du sein.
Comme perspectives futures, il serait intéressant d'explorer des architectures plus récentes, d'utiliser des bases de données plus importantes et d'intégrer des techniques d'augmentation de données afin d'améliorer davantage les performances du système. L'intégration d'un système complet de détection, de localisation et de classification pourrait également constituer une étape importante vers une aide au diagnostic médical plus performante.