
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université d'Ain-Temouchent Belhadj Bouchaib – UATBB-
Faculté des sciences exactes
Département Mathématique et Informatique



MÉMOIRE

Présenté en vue de l'obtention du diplôme de Master

Domaine : Mathématique et Informatique

Filière : Informatique

Spécialité : Réseau et ingénierie des données

Présenté Par :

M.DIAKITE MADIASSA LAKAMY

THEME

Développement d'une application mobile d'identification des
plantes médicinales.

Soutenu le 24/06/2024

Devant le jury composé de :

| | | |
|---|---------|-----------|
| Présidente : BELGRANA FATIMA ZOHRA | « MCA » | U.B.B.A.T |
| Examineur : MEDEDJEL MANSOUR | « MCA » | U.B.B.A.T |
| Encadrante : BEDDAD FATIMA | « MCB » | U.B.B.A.T |
| CO-Encadrante : BOUSAHBA NASSIMA | « MCB » | U.H.B.C |

Année universitaire : 2023-2024

Remerciements

Nous exprimons notre gratitude à ALLAH le tout puissant pour nous avoir accordé la sante et la volonté de mener à bien ce projet, El hamdollah !

Tout d'abord, nous remercions chaleureusement notre Encadrante, Madame BEDAD FATIMA, pour son orientation, sa confiance en nous et son expertise dans le domaine. Grâce à ses encouragements et ses critiques constructives, nous avons pu repousser nos limites et trouver des solutions novatrices.

Nous tenons à exprimer notre gratitude envers notre Co-encadrante, NASSIMA BOUSSAHABA enseignante à l'université de Chlef, pour sa disponibilité et ses précieux conseils qui ont grandement contribué à nos avancées. Nous le remercions également pour sa patience et son soutien constant.

Nos remerciements vont également aux membres du jury qui ont accepté d'évaluer notre travail et de l'enrichir de leurs connaissances et suggestions.

Nous tenons à exprimer notre reconnaissance envers le corps professoral et administratif de l'université Belhadj Bouchaib d'Ain Temouchent, qui a contribué à notre réussite universitaire.

Nous tenons à remercier tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce projet.

Merci du fond du cœur.

Dédicaces

J'ai le grand plaisir de dédier ce modeste travail :

A ma très chère mère :

Celle qui a veillé à mes cotés depuis mon jeune âge jusqu'à aujourd'hui tu n'es pas seulement une mère pour moi mais une meilleure amie, ma sœur affectueuse et mon amie fidèle. Je la trouve toujours à mes côtés, celle qui est la sérénité de mon cœur, mon réconfort et mon plaisir dans la vie. Sa présence est ma source d'inspiration et ma force pour affronter les différents obstacles.

A mon très cher père :

Tu as été toujours mon soutien et ma sécurité dans la vie, mon courage, ma confiance et ma force. Tu cherches toujours mon confort et que rien ne m'empêche de terminer mes études. Tu as toujours été à mes côtés pour me soutenir et m'encourager. Que ce travail reflète ma gratitude et mon affection.

A mes sœurs :

Puisse Dieu vous donnez la santé, bonheur, courage et surtout réussite.

A tous mes meilleurs amis.

Et tous ceux qui m'ont aidé et contribué à ce modeste travail.

Je dédie ce fruit de plusieurs années d'études en signe de reconnaissance.

Résumé

Notre projet vise à développer une Application mobile fonctionnant sous un système de vision par ordinateur avec objectif principal d'identifier les plantes médicinales. L'objectif du système est de permettre une identification précise des plantes grâce à l'utilisation de méthodologies sophistiquées de vision par ordinateur, en particulier la segmentation des instances. Le système utilise une approche de segmentation d'instance basée sur YOLOv8, qui est un modèle d'apprentissage profond compétent et résilient qui a la capacité de détecter et de localiser avec précision des objets dans une image. Grâce à la mise en œuvre de ce modèle dans le cadre de l'identification des plantes, l'application mobile est capable de discerner et d'identifier efficacement les plantes. La méthodologie de segmentation des instances utilisée dans ce système discerne efficacement les plantes facilitant ainsi une identification efficace des plantes en utilisant les feuilles. L'intégration de YOLOv8 permet au système d'améliorer son efficacité et sa fiabilité dans l'évaluation de l'identification des plantes, facilitant ainsi une détection efficace.

Mots-clés : Plantes médicinales, Intelligence artificielle, Vision par ordinateur, Segmentation d'instance, Application mobile.

ملخص

يهدف مشروعنا إلى تطوير تطبيق جوال يعمل بنظام رؤية حاسوبية بهدف رئيسي هو تحديد النباتات الطبية. الهدف من النظام هو تمكين التحديد الدقيق للنبات من خلال استخدام منهجيات الرؤية الحاسوبية المتطورة، وخاصة تجزئة الحالة. يستخدم النظام نهج تجزئة المثيلات استنادًا إلى YOLOv8 ، وهو نموذج تعلم عميق يتسم بالكفاءة والمرونة ولديه القدرة على اكتشاف الكائنات في الصورة وتحديد موقعها بدقة. ومن خلال تنفيذ هذا النموذج في تحديد النباتات، يصبح تطبيق الهاتف المحمول قادرًا على تمييز النباتات وتحديد شكل فعال. تعمل منهجية تجزئة المثيلات المستخدمة في هذا النظام على تمييز النباتات بشكل فعال وبالتالي تسهيل التعرف الفعال على النباتات باستخدام الأوراق. يتيح تكامل YOLOv8 للنظام تحسين كفاءته وموثوقيته في تقييم تحديد هوية النبات، وبالتالي تسهيل الاكتشاف الفعال

الكلمات المفتاحية: النباتات الطبية, الذكاء الاصطناعي, الرؤية الحاسوبية, تجزئة المثيلات, تطبيق الجوال

Abstract

Our project aims to develop a Mobile Application operating under a computer vision system with the main objective of identifying medicinal plants. The aim of the system is to enable accurate plant identification through the use of sophisticated computer vision methodologies, particularly instance segmentation. The system uses an instance segmentation approach based on YOLOv8, which is a competent and resilient deep learning model that has the ability to accurately detect and localize objects in an image. By implementing this model in plant identification, the mobile application is able to discern and identify plants effectively. The instance segmentation methodology used in this system effectively discerns plants thereby facilitating effective plant identification using leaves. The integration of YOLOv8 allows the system to improve its efficiency and reliability in evaluating plant identification, thereby facilitating effective detection.

Keywords : Medicinal plants, Artificial intelligence, Computer vision, Instance segmentation, Mobile application.

Sommaire

| | |
|--|-----------|
| Liste des figures | 8 |
| Liste des tableaux | 10 |
| Abréviation | 11 |
| Introduction générale | 13 |
| Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales | 15 |
| 1.1 Introduction | 16 |
| 1.2 Intelligence artificielle | 16 |
| 1.3 Apprentissage automatique | 17 |
| 1.3.1 SVM | 18 |
| 1.3.2 KNN | 19 |
| 1.3.3 ANN | 19 |
| 1.4 Apprentissage profond | 20 |
| 1.4.1 CNN | 21 |
| 1.4.2 PNN | 23 |
| 1.5 Etat de l'art : | 23 |
| 1.5.1 Première partie : méthodes d'apprentissage profond : | 23 |
| 1.5.2 Deuxième partie : méthodes d'apprentissage automatique : | 24 |
| 1.6 Identification des plantes à l'aide des applications mobile. | 25 |
| 1.7 Conclusion | 28 |
| | |
| Chapitre 2 : Vision par ordinateur and segmentation d'instance | 29 |
| 2.1 Introduction | 30 |
| 2.2 Vision par ordinateur. | 30 |
| 2.2.1 Les quatre taches courantes en vison par ordinateur | 30 |
| 2.2.2 Utilisation de la vision par ordinateur | 31 |
| 2.3 Détection d'objet | 32 |
| 2.3.1 Techniques de détection d'objets | 32 |
| 2.4 Segmentation d'image | 34 |
| 2.4.1 Les méthodes de segmentation d'image | 35 |
| 2.4.2 Segmentation d'image en utilisant l'apprentissage profond | 35 |
| 2.4.3 Types de taches de la segmentation d'images | 39 |
| 2.4.4 Avantages et inconvénients pour chaque type de segmentation d'image | 40 |

| | |
|---|-----------|
| 2.4.5 Pourquoi la Segmentation d'instance ? | 42 |
| 2.5 Conclusion | 43 |
| Chapitre 3 : Conception & Réalisation | 44 |
| 3.1 Introduction | 45 |
| 3.2 Architecture du système | 45 |
| 3.3 Collecte et prétraitement des données | 46 |
| 3.3.1 Annotation des données | 47 |
| 3.3.2 Augmentation des données | 47 |
| 3.4 YOLO | 48 |
| 3.4.1 YOLO architecture | 49 |
| 3.4.2 YOLO versions | 49 |
| 3.4.3 YOLOv8 | 49 |
| 3.4.4 Pourquoi devons nous utiliser YOLOv8 ? | 50 |
| 3.5 YOLO pour la segmentation d'instance | 50 |
| 3.6 Segmentation YOLOv8 pour la détection des plantes médicinales | 50 |
| 3.7 Intégration de la segmentation YOLOv8 dans l'application mobile | 51 |
| 3.8 Conclusion | 52 |
| Chapitre 4 : Implémentation & Résultats | 53 |
| 4.1 Introduction | 54 |
| 4.2 Conception logiciel | 54 |
| 4.3 Matériel | 55 |
| 4.4 Interface utilisateur | 56 |
| 4.5 Comment fonctionne l'identification des images dans l'application | 56 |
| 4.6 Heuristique utilisée dans la conception | 57 |
| 4.7 Création de notre modèle d'apprentissage profond | 58 |
| 4.7.1 Affiner le modèle YOLOv8 pré-entraîne | 58 |
| 4.7.2 Entraîner le modèle | 59 |
| 4.8 Résultat et discussion | 60 |
| 4.8.1 La moyenne de la précision moyenne | 60 |
| 4.8.2 Affichage des courbes | 60 |
| 4.8.3 Affichage de la matrice de confusion | 62 |
| 4.8.4 Résultats | 64 |

| | |
|---|----|
| 4.8.5 Evaluation du modèle | 66 |
| 4.9 Comparaison avec YOLOv5 et YOLOv7 | 67 |
| 4.10 Conclusion | 68 |
| Conclusion générale | 69 |
| Travaux futurs | 70 |
| Glossaire | 71 |
| Bibliographies | 72 |

Liste des figures

Liste des figures

| | | |
|------|---|----|
| 1.1 | Les sous domaines de l'intelligence artificielle [11]. | 18 |
| 1.2 | Architecture de l'algorithme SVM.[15] | 19 |
| 1.3 | Algorithme KNN.[61] | 20 |
| 1.4 | Architecture du ANN [33] | 21 |
| 1.5 | Les performances de l'apprentissage en profondeur.[66] | 22 |
| 1.6 | Architecture du CNN.[70] | 23 |
| 1.7 | Architecture du PNN.[12] | 24 |
| 1.8 | Logo de l'application PictureThis... | 27 |
| 1.9 | Logo de l'application PlantSnap | 27 |
| 1.10 | Logo de l'application NatureID | 27 |
| 1.11 | Logo de l'application PlantNet. | 28 |
| 1.12 | Logo de l'application FloraIncognita... | 28 |
| 2.1 | Domaine d'application de la vision par ordinateur.[67] | 32 |
| 2.2 | Exemple de détection d'objet | 33 |
| 2.3 | Classification des techniques de détection d'objet. | 34 |
| 2.4 | Méthode de segmentation d'image.[48] | 36 |
| 2.5 | CNN Architecture pour la Segmentation sémantique.[17] | 37 |
| 2.6 | Segmentation sémantique déconvolutive suivie d'un réseau convolutionnel de couche VGG-16.[46] | 38 |
| 2.7 | U-Net et V-Net.[14] | 38 |
| 2.8 | Réseaux multi-échelles et pyramidaux.[45] | 39 |
| 2.9 | R-CNN.[20] | 40 |
| 2.10 | Les types de segmentation d'images.[71] | 40 |
| 3.1 | Architecture du système | 46 |
| 3.2 | Annotation avec outils Roboflow. | 48 |
| 3.3 | Echantillon d'augmentation. | 49 |
| 3.4 | YOLO Architecture.[76] | 50 |
| 3.5 | YOLO <u>evolution</u> .[64] | 50 |
| 3.6 | Les étapes de déploiement sur votre application. | 53 |
| 4.1 | Vitrine des pages d'intégration. | 58 |
| 4.2 | Design des pages principales. | 59 |
| 4.3 | Tutoriel de l'application. | 60 |

Liste des figures

| | | |
|------|---|----|
| 4.4 | Data.yamll | 61 |
| 4.5 | Pourcentages des données d'entraînement, de validation et de test | 62 |
| 4.6 | BoxF1-curve. | 63 |
| 4.7 | Box P-curve. | 63 |
| 4.8 | Box PR-curve | 63 |
| 4.9 | Box R-curve | 63 |
| 4.10 | Mask F1-curve | 63 |
| 4.11 | Mask P-curve | 63 |
| 4.12 | Mask PR-curve | 64 |
| 4.13 | Mask R-curve | 64 |
| 4.14 | La matrice de confusion. | 65 |
| 4.15 | La fonction de perte change au fil de la formation. | 66 |
| 4.16 | La fonction de perte change au fil de la validation. | 67 |
| 4.17 | Résultat de la précision et du rappel. | 67 |
| 4.18 | Résultat mAP50-95. | 68 |
| 4.19 | Prédictions du modèle sur l'ensemble de données de test. | 69 |

Liste des tableaux

Liste des tableaux

| | | |
|-----|--|----|
| 1.1 | Résumé des travaux déjà effectués en Apprentissage Profond | 24 |
| 1.2 | Résumé des travaux déjà effectués en Apprentissage automatique | 25 |
| 2.1 | Avantage et Inconvénient de la segmentation d'instance. | 40 |
| 2.2 | Avantage et Inconvénient de la Segmentation Sémantique. | 41 |
| 2.3 | Avantage et Inconvénient de la Segmentation Panoptique. | 42 |
| 3.1 | Images de quelques plantes. | 46 |
| 4.1 | Comparaison avec d'autre modele | 67 |

Liste des Acronymes

Liste des Acronymes

- AA Apprentissage Automatique
- ANN Artificial neural network
- AP Apprentissage profond
- CNN Convolutional Neural Network
- CPCA Contrastive Principal
Component Analysis
- DPM Deformable Parts Mode
- HOG Histogramme de Gradient
Orienté
- IA Intelligence artificielle
- KNN K-Nearest Neighbors
- LDR Logistic Decision Regression
- NLP Le traitement du langage
naturel
- PNN Probabilistic Neural Network
- RCNN Region-based Convolutional
Neural Networks
- SPP Spatial Pyramid Pooling
Network
- SSD Single Shot Detector
- SVM Support vector machine
- VJ Viola Jones
- YOLO You Only Look Once

Introduction générale

Introduction générale

Introduction générale

Les plantes thérapeutiques jouent un rôle essentiel dans la vie des populations rurales, notamment, où les savoirs ancestraux sont transmis de génération en génération. Toutefois, ces savoirs sont fréquemment menacés de disparition sans être transmis ou transcrits.

Afin de sauvegarder et mettre en valeur ces connaissances, il est primordial de mettre au point des outils contemporains pour recueillir et analyser les informations sur les plantes médicinales. Pour la médecine traditionnelle, il est crucial de reconnaître les plantes médicinales afin de distinguer les différentes espèces et de garantir l'efficacité et la sécurité des traitements à base de plantes.

On peut identifier les plantes médicinales en utilisant des indices quantitatifs grâce à une application mobile. Grâce à cette application, il serait possible de collecter des renseignements précis sur les plantes médicinales, tels que leurs propriétés, leurs utilisations traditionnelles etc...

L'identification des plantes médicinales innovante et efficace est réalisée en utilisant les méthodes de vision par ordinateur et d'apprentissage automatique. On utilise la technologie de détection d'objets, notamment YOLO, afin de repérer les plantes. Le traitement en temps réel des images de feuilles par l'algorithme YOLO permet d'obtenir un taux d'images de 45 images par seconde, ce qui dépasse les performances des méthodes de détection d'objets existantes.

L'utilisation de la segmentation des instances constitue une approche puissante pour l'identification et la différenciation précises d'instances individuelles des plantes, permettant ainsi la détection précise des plantes. Grâce à l'utilisation d'algorithmes sophistiqués de vision par ordinateur, la segmentation des instances surpasse les techniques conventionnelles en fournissant un examen plus complet des plantes médicinales.

Ce mémoire est organisé en quatre parties principales, chacune traitant un aspect spécifique du projet. Un aperçu de ces parties est présenté comme suit :

Le premier chapitre procédera à une analyse complète des principes de base de l'intelligence artificielle (IA), en Apprentissage Profond et en Apprentissage Automatique, en mettant l'accent sur l'exploration des avancées actuelles dans ces deux dernières méthodologies utilisées aux fins de détecter et segmenter les plantes médicinales.

Introduction générale

Dans le deuxième chapitre, nous examinerons le domaine de la vision par ordinateur, en mettant l'accent sur la segmentation des instances.

Les données utilisées pour la détection des plantes, y compris leur collecte, seront présentées dans le chapitre 3. En outre, ce chapitre exposera le modèle YOLO et ces différentes versions et exposera les étapes nécessaires pour le déploiement de ce modèle.

Le chapitre final présentera les informations essentielles concernant les logiciels et le matériel utilisés dans le cadre de ce projet, ainsi que les étapes nécessaires pour développer le modèle YOLOv8, suivies d'une discussion sur les résultats obtenus.

Et pour la conclusion récapitule tous les résultats obtenus, évalue les objectifs atteints et suggère des améliorations futures.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

1.1 Introduction

Les végétaux jouent un rôle crucial dans l'écosystème en tant que producteurs pour les êtres humains et les autres organismes vivants.

Depuis des siècles, l'homme utilise différentes plantes dans sa médecine traditionnelle. La plupart des plantes sont perçues comme étant les ressources indispensables pour le bien-être des individus, car elles ont un impact important sur la formation et le développement de la civilisation humaine. Les plantes médicinales ont été employées pendant des milliers d'années afin de préserver la santé publique.

Dans cette optique, les chercheurs ont mis au point une approche automatisée qui aide considérablement les herboristes à détecter et identifier des plantes en fonction de leur apparence et de leurs caractéristiques visuelles.

Pour obtenir les meilleurs résultats, les chercheurs ont utilisé diverses approches, y compris l'apprentissage automatique et l'apprentissage profond, pour créer ces systèmes. Nous ne pouvons cependant pas décrire ces procédures sans une compréhension de base. Tous les détails seront inclus dans ce chapitre.

1.2 Intelligence artificielle

L'intelligence artificielle (IA) est une technique qui permet à des ordinateurs ou à des machines d'imiter la capacité humaine à agir, se comporter ou fonctionner. L'IA peut apprendre des données passées et est capable de reconnaître des modèles et des comportements. Cela permet à une machine d'apprendre de l'expérience et de tolérer son comportement [18].

La figure 1.1 présente les sous domaines de l'intelligence artificielle.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

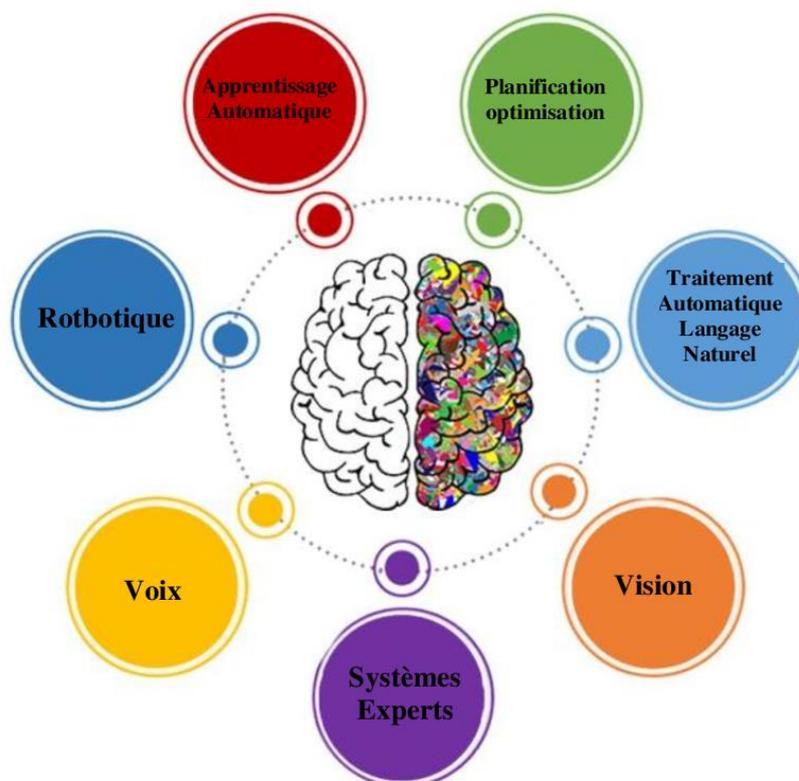


Figure 1.1 : Les sous domaines de l'intelligence artificielle [11]

1.3 Apprentissage automatique

L'apprentissage automatique (AA) n'est pas un terme nouveau dans le domaine informatique. Cependant, il s'agit d'un terme en pleine expansion et en pleine évolution. L'AA est un sujet qui étudie comment utiliser des machines pour reproduire les capacités et le comportement humains. Elle est considérée comme une technologie utilisée pour répondre aux questions et tirer des significations des données. Il rassemble les statistiques et l'informatique pour permettre aux ordinateurs ou aux machines d'effectuer une tâche donnée [18].

L'apprentissage automatique a démontré des résultats prometteurs dans la construction de modèles de classification dans divers domaines, notamment la vision par ordinateur, la reconnaissance vocale, reconnaissance de texte et jeux d'IA [24].

On distingue plusieurs types d'apprentissage automatique, comme l'apprentissage supervisé, où on peut trouver des algorithmes tels que SVM, KNN, et l'apprentissage non supervisé, il existe différentes options telles que Kmeans et l'apprentissage par renforcement, où on peut trouver des algorithmes tels que le Q-Learning et les méthodes de gradients de politique.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

Quelques algorithmes utilisés fréquemment dans les études de détection des plantes seront présentés.

1.3.1 SVM

Support Vector Machine est l'une des techniques classiques d'apprentissage automatique qui peut encore aider à résoudre les problèmes de classification du Big Data [38].

Comment elle fonctionne ?

L'algorithme de machine à vecteurs de support (SVM) peut classer les données linéaires et non linéaires. Il mappe d'abord chaque élément de données dans l'espace de fonctionnalités à n dimensions où correspond le nombre de fonctionnalités. Il définit ensuite le plan atmosphérique qui sépare les éléments de données en deux parties tout en maximisant la distance marginale pour les deux classes et en minimisant les erreurs de classification. La distance marginale d'une classe est la distance entre l'hyperplan de décision et sa position la plus proche qui est membre de cette classe [37].

La figure 1.2 présente l'architecture du modèle SVM.

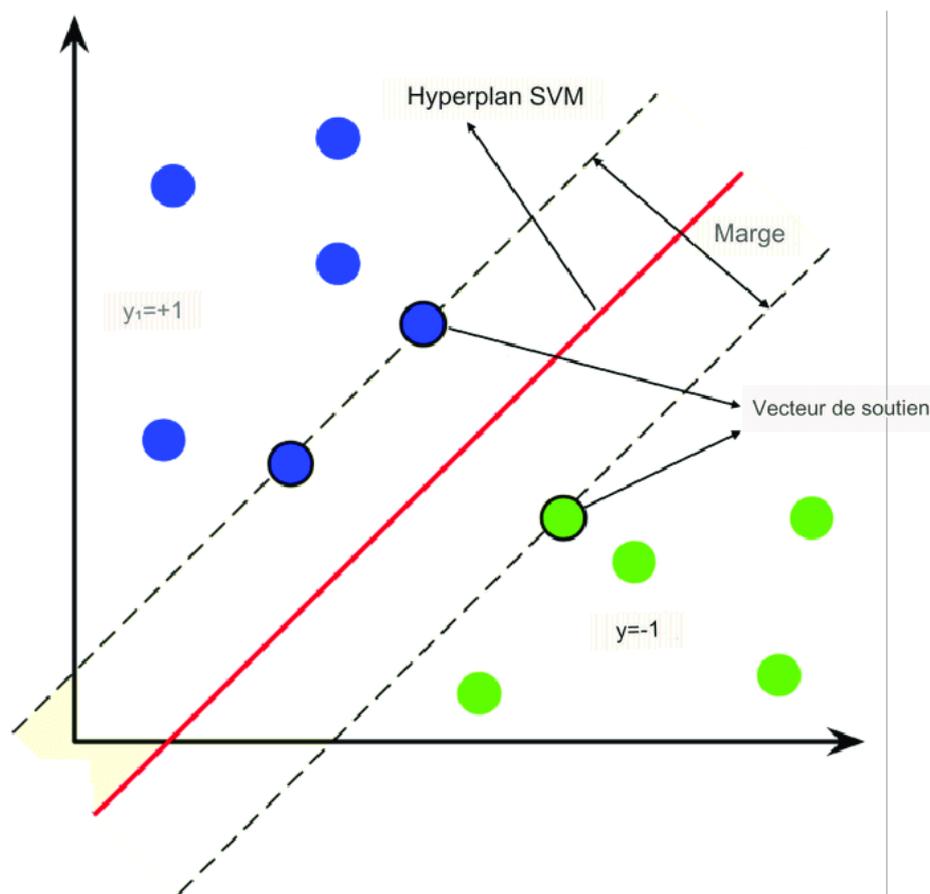


Figure 1.2: Architecture de l'algorithme SVM [15]

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

1.3.2 KNN

C'est un algorithme de classification non paramétrique, c'est-à-dire qu'il ne fait aucune hypothèse sur l'ensemble de données brutes. Il s'agit d'un algorithme d'apprentissage supervisé.

Un ensemble de données de formation étiquetées est fourni dans lequel les points de données sont classés en différentes catégories, de sorte que la catégorie de données non étiquetée puisse être prédite [23].

La figure 1.3 présentera la visualisation sur comment travail l'algorithme du KNN.

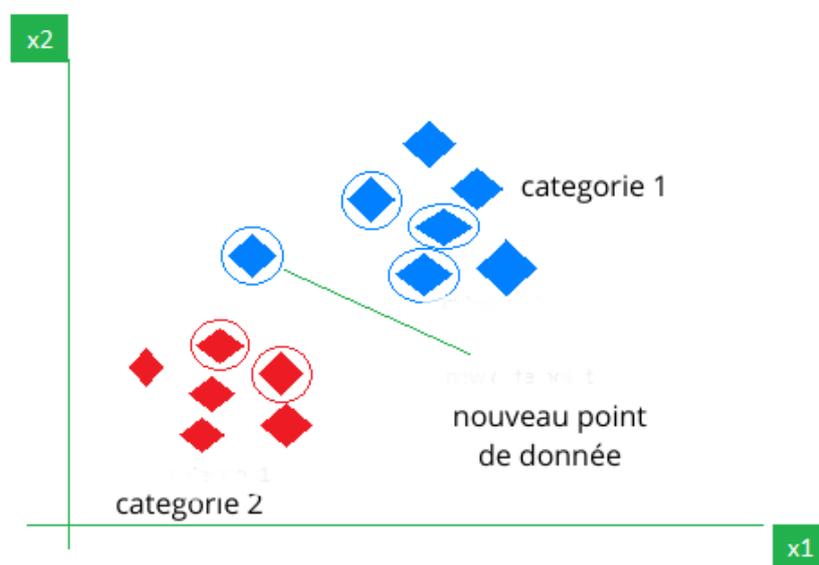


Figure 1.3: Algorithme KNN [61]

1.3.3 ANN

Les réseaux de neurones artificiels sont une méthode de calcul qui construit plusieurs unités de traitement basées sur des connexions interconnectées. Le réseau est constitué d'un nombre arbitraire de cellules, de nœuds, d'unités ou de neurones qui connectent l'ensemble d'entrées à la sortie. Il fait partie d'un système informatique qui imite la façon dont le cerveau humain analyse et traite les données [36].

La figure 1.4 présente l'architecture du ANN.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

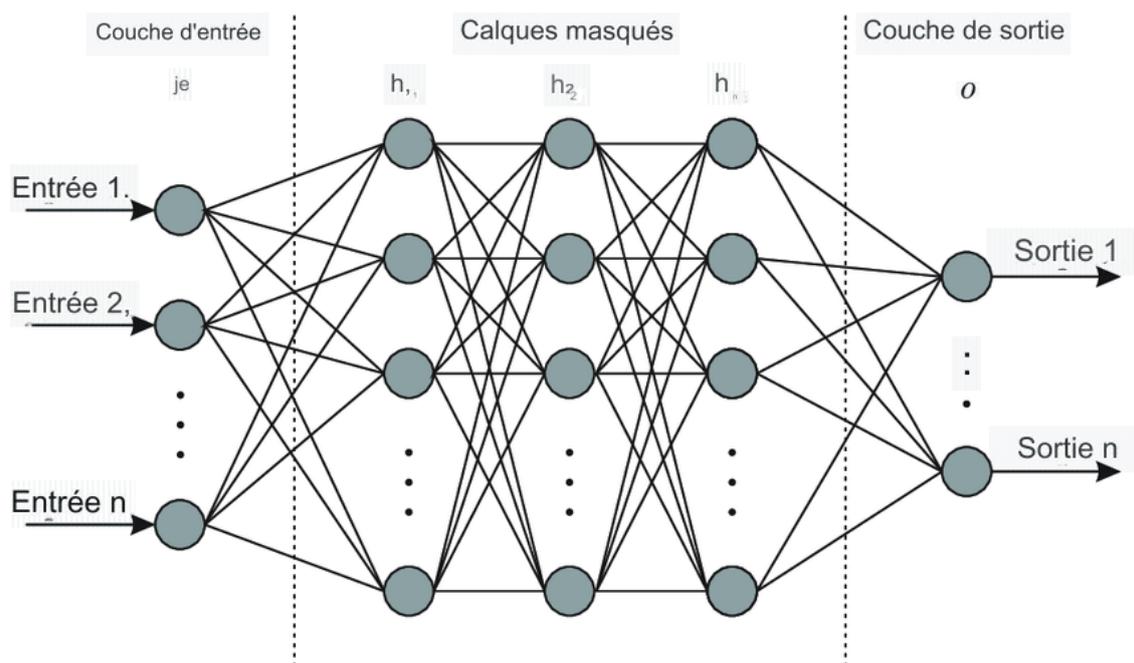


Figure 1.4: Architecture du ANN [33]

La figure 1.4 montre l'architecture générale de l'ANN, qui comporte une couche d'entrée, un ensemble de couches cachées et une couche de sortie. Dans chaque couche cachée et de sortie, il y a des neurones artificiels interconnectés via des poids adaptatifs. Ces poids sont calibrés via un processus de formation avec des données d'entrée-sortie. Pour chaque neurone artificiel, il existe une fonction d'activation, qui peut être n'importe quelle fonction d'intervalle $[-1,1]$; les fonctions d'activation les plus courantes sont la sigmoïde tangente et la sigmoïde logarithmique [33].

1.4 Apprentissage profond

L'apprentissage profond (AP) est un sous-domaine de l'apprentissage automatique qui utilise des réseaux de neurones artificiels pour apprendre et résoudre des problèmes complexes. Il combine l'intelligence artificielle et l'apprentissage automatique.

L'apprentissage profond repose sur un réseau de neurones artificiels (ANN) multicouche. Le concept est basé sur les réseaux neuronaux biologiques chez l'homme. Ces réseaux contiennent des nœuds et des multicouches qui communiquent entre eux pour comprendre et analyser les données reçues. AP se compose de plusieurs couches, tandis que chaque couche peut être constituée de plusieurs neurones.

Le neurone de chaque couche est appelé unité de traitement. Le rôle principal des neurones est de multiplier les valeurs saisies avec les poids donnés et calculer la somme des

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

résultats. C'est ainsi que les modèles d'apprentissage profond acquièrent une plus grande précision [18]. Ou on peut trouver quelques applications de l'apprentissage profond : Vision par ordinateur pour la Reconnaissance d'images, segmentation sémantique, détection d'objets, traitement du langage naturel pour la traduction automatique, analyse de sentiment, génération de texte, robotique pour le contrôle de robots, navigation autonome, prise de décision intelligente.

La figure 1.5 présente les performances de l'apprentissage profond.

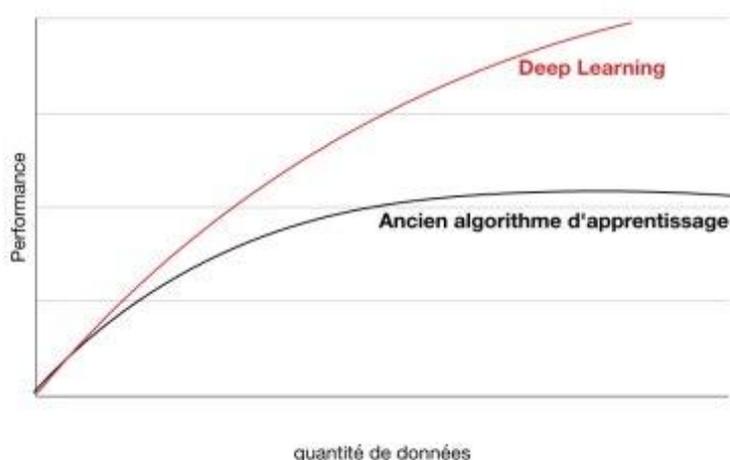


Figure 1.5: Les performances de l'apprentissage en profondeur [66]

De nombreuses équipes scientifiques bien constituées ont utilisé et expérimenté une vaste gamme de techniques et d'ensembles de données en apprentissage profond. Quelques algorithmes utilisés fréquemment dans les études de détection des plantes en utilisant l'apprentissage profond seront présentés.

1.4.1 CNN

Les réseaux de neurones convolutifs (CNN) sont une forme de réseau de neurones artificiels qui excelle dans le traitement des entrées de pixels. Ils sont souvent utilisés dans le traitement et la reconnaissance d'images [63].

Les CNN sont des systèmes d'intelligence artificielle (IA) basés sur le traitement d'images qui utilisent l'apprentissage en profondeur pour effectuer des tâches à la fois

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

productives et descriptives, qui incluent souvent la vision industrielle, la reconnaissance d'images et de vidéos, les systèmes de recommandation et l'analyse naturelle et le traitement du langage (NLP).

La figure 1.6 présente un exemple du CNN.

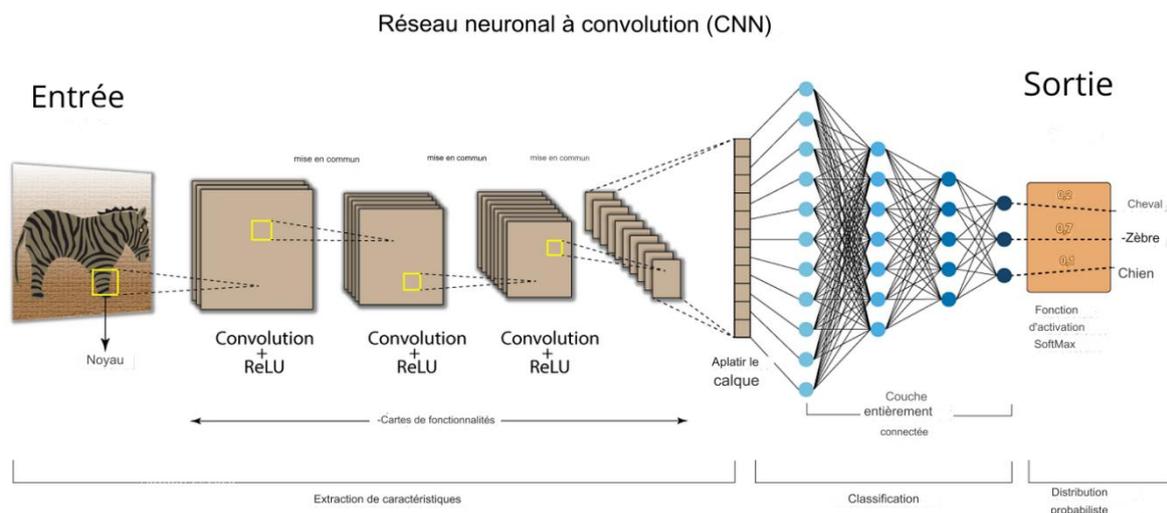


Figure 1.6: Architecture du CNN [70]

Les composants de l'Architecture du CNN sont :

- **Bloc de traitement** : Un bloc de traitement est composé de plusieurs couches :
 - *Couche de convolution (CONV)* : Utilisez des filtres convolutifs sur les images afin de repérer des caractéristiques locales.
 - *Couche de correction (ReLU)* : Utilisez la fonction d'activation ReLU pour favoriser l'apprentissage plus rapide et efficace.
 - *Couche de pooling (POOL)* : Réduit la taille de l'image intermédiaire par sous-échantillonnage non linéaire.
- **Couches intermédiaires** : Les blocs de traitement s'enchaînent jusqu'à des couches finales qui réalisent la classification de l'image et le calcul de l'erreur entre la prédiction et la valeur cible:
 - *Couche entièrement connectée (FC)* : Effectue une classification finale en générant un vecteur de probabilités.
 - *Couche de perte (LOSS)* : Calcule l'erreur entre la prédiction et la valeur cible.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

1.4.2 PNN

Les réseaux de neurones probabilistes (PNN) sont un type de réseau de neurones artificiels à action directe qui est étroitement lié à l'estimation de la densité du noyau via la fenêtre de Parzen qui se rapproche asymptotiquement de la minimisation optimale des risques de Bayes. Cette technique est largement utilisée pour estimer les densités conditionnelles de classe (également appelées vraisemblance) dans les tâches d'apprentissage automatique telles que l'apprentissage supervisé.[69]

La figure 1.7 présente l'architecture du PNN.

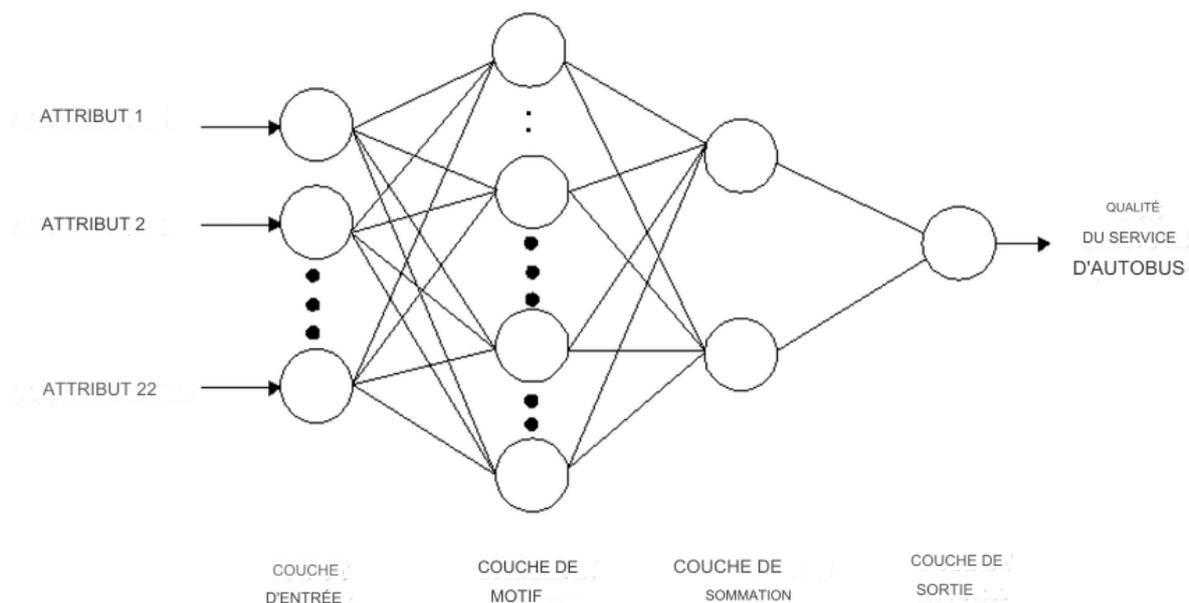


Figure 1.7: Architecture du PNN [12]

1.5 Etat de l'art

Le développement d'une application mobile d'identification de plantes médicinales à l'aide d'indices quantitatifs nécessite une combinaison de technologies telles que l'apprentissage automatique, et en apprentissage profond.

1.5.1 Première partie : méthodes d'apprentissage profond

Le tableau 1.1 présentera les principales recherches réalisées en utilisant l'apprentissage profond.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

| Apprentissage profond | | | | | |
|-----------------------|------------------------------|----------------------------|-------|-----------------------|----------|
| N | Auteur | Nom de la méthode | Année | Base de données | Résultat |
| 1 | Stephen Gang Wu [42] | PNN | 2007 | Leaf dataset | 90%. |
| 2 | Mostafa Mehdipour Ghazi [26] | GoogLeNet, AlexNet, VGGNet | 2017 | LifeCLEF 2015 | 80%. |
| 3 | Yu Sun [53] | ResNet 26 | 2017 | BJFU100 | 91,78%. |
| 4 | Md. Khairul Islam [25] | Yolov2 plus CNN | 2019 | PASCAL VOC Dataset | 96%. |
| 5 | Wan Muhammad Arief [47] | CNN | 2022 | Leaf dataset malaisie | 99 %. |
| 6 | RahimAzadnia[3] | CNN | 2022 | Leaf dataset | 99,3% |
| 7 | Owais A. Malik [31] | EfficientNet-B1 | 2022 | Leaf dataset | 87%. |
| 8 | Youssef Sharrab [52] | VCG-16 | 2023 | ImageNet | 89%. |
| 9 | Yang, T [44] | DEEPLABV3+, YOLOv8 | 2023 | Dataset jekyll | 90,8 %. |

Table 1.1: Résumé des travaux déjà effectués en Apprentissage Profond

1.5.2 Deuxième partie : méthodes d'apprentissage automatique :

Le tableau 1.2 présentera les principales recherches réalisées en utilisant l'apprentissage automatique.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

| Apprentissage Automatique | | | | | |
|---------------------------|------------------------|---------------------------|-------|--------------------------|----------|
| N | Auteur | Nom de la méthode | Année | Base de données | Résultat |
| 1 | Robert G. de Luna [35] | ANN | 2017 | Leaf image | 98,61%. |
| 2 | Agus Ambarwari [2] | SVM | 2020 | Flavia dataset | 83%. |
| 3 | Shubham Mahajan [40] | SVM, AdaboostT, KNN | 2021 | Flavia dataset | 95,85%. |
| 4 | Kannan Kattari [22] | CPCA | 2023 | Medicinal plants dataset | 79 %. |
| 5 | Kannan Kattari[22] | LDR | 2023 | Medicinal plants dataset | 90 %. |
| 6 | Wirdayanti [50] | KNN | 2023 | Herbal Leaf | 78,67%. |

Table 1.2: Résumé des travaux déjà effectués en Apprentissage automatique

1.6 Identification des plantes à l'aide des applications mobiles

L'identification des plantes peut être effectuée sans effort grâce à l'utilisation d'applications mobiles. Ces applications utilisent une technologie de reconnaissance d'images, permettant aux utilisateurs de simplement prendre une photo d'une plante inconnue et de recevoir des détails complets concernant son identification, ses caractéristiques distinctives et ses applications possibles. Dans la section suivante, nous présenterons un résumé des principales applications mobiles actuellement disponibles pour l'identification des plantes.

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

a. **PictureThis: Une Porte d'Entrée Vers le Monde Végétal**



Figure 1.8: Logo de l'application PictureThis ¹.

PictureThis (Figure 1.8) est une application mobile populaire qui utilise l'intelligence artificielle pour identifier avec précision les plantes à partir d'une simple photo. L'application est capable de reconnaître une grande variété de plantes, des fleurs sauvages aux plantes d'intérieur.

b. **PlantSnap: Capturer et Apprendre**



Figure 1.9: Logo de l'application PlantSnap ².

PlantSnap (Figure 1.9) est une autre application d'identification de plantes très performante. Utilisant un algorithme de reconnaissance d'images avancé, l'application est capable d'identifier plus de nombreuses espèces de plantes à partir d'une simple photo.

c. **NatureID: Une Approche Pédagogique Identification Précise**



Figure 1.10: Logo de l'application NatureID ³

¹<https://www.picturethisai.com/fr/>

² <https://www.plantsnap.com/>

³ <https://natureid-plant-identification.fr.softonic.com/android>

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

Grâce à sa base de données étendue de plus de 5 000 espèces de plantes, NatureID (Figure 1.10) offre une identification précise des plantes à partir de photos.

d. Pl@ntNet: Une Communauté d'Observateurs



Figure 1.11: Logo de l'application PlantNet⁴

Pl@ntNet (Figure 1.11) est une application interactive qui favorise un sentiment de communauté en invitant les utilisateurs à partager leurs observations de plantes et à participer activement à la création d'une base de données complète. En tirant parti des contributions des utilisateurs, Pl@ntNet affine continuellement son algorithme d'identification, ce qui donne lieu à une application de plus en plus précise au fil du temps.

e. FloraIncognita: Une Approche Scientifique



Figure 1.12: Logo de l'application FloraIncognita⁵.

Grâce à son algorithme de reconnaissance d'images de pointe basé sur l'apprentissage profond, FloraIncognita (Figure 1.12) est capable d'identifier avec précision diverses espèces végétales. L'application est équipée d'une vaste base de données contenant des millions d'images de plantes, garantissant des résultats précis.

⁴ <https://plantnet.org/>

⁵ <https://floraincognita.com/>

Chapitre 1 : Etat de l'art sur l'identification des plantes médicinales

1.7 Conclusion

Le présent chapitre a fourni un résumé élaboré des progrès actuels dans le domaine de la détection des plantes et quelques applications déjà existantes dans l'identification des plantes. Le chapitre suivant se penchera sur les diverses méthodologies et stratégies qui ont émergé dans ces dernières années pour résoudre ce problème important.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

Chapitre 2 : Vision par ordinateur et segmentation d'instance

2.1 Introduction

La vision par ordinateur et la segmentation des instances sont essentielles pour détecter les plantes. À l'aide d'algorithmes et de modèles d'apprentissage automatique, ces nouvelles technologies analysent le numérique des images de cultures pour identifier et séparer correctement les composants botaniques spécifiques.

Traditionnelles les botanistes utilisaient la biologie moléculaire et les caractéristiques cellulaires des feuilles pour détecter les plantes. Cependant, l'adoption de telles approches pour l'identification des plantes souffre de graves désavantages. Premièrement, ces approches nécessitent effectuer plusieurs étapes critiques. Deuxièmement, de telles approches sont tellement fastidieuses et prennent beaucoup du temps. D'autre part, la segmentation des instances et la vision par ordinateur offrent une solution pratique pour ces problèmes.

2.2 Vision par ordinateur

La vision par ordinateur [21] est un domaine de l'intelligence artificielle qui entraîne les ordinateurs à interpréter et à comprendre le monde visuel pour une variété de tâches passionnantes en aval telles que les voitures autonomes, les achats sans caisse, les villes intelligentes, la détection du cancer et en savoir plus. Le domaine de la vision par ordinateur a été révolutionné par l'apprentissage profond au cours de la dernière décennie.

2.2.1 Les quatre tâches courantes en Computer Vision

Les principales tâches de la vision par ordinateur [6] sont la classification d'images, la détection d'objets, la segmentation sémantique et la segmentation d'instances.

- **Classification** : c'est lorsqu'une image se voit attribuer une seule étiquette représentant l'objet le plus important contenu dans une image d'entrée. En plus de la classification simple, la localisation englobe également le dessin d'un cadre de délimitation sur une section de l'image d'entrée d'origine. Par conséquent, englobant strictement le seul objet classé par le modèle.
- **Détection d'objets** : consiste à classer plusieurs objets à partir d'une seule image d'entrée. Cette détection se fait dans le style de la localisation. La sortie est alors composée de plusieurs cadres de délimitation et étiquettes correspondant aux objets détectés.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

- **Segmentation d'instance** : est similaire à la détection d'objets, mais avec quelques ajouts. Au lieu d'avoir des cadres de délimitation, la sortie est désormais un masque de pixels sur l'image. Plus important encore, cela ajoute la possibilité de distinguer les objets de la même classe.
- **Segmentation sémantique** : fait référence au partitionnement d'une image d'entrée donnée en régions de pixels qui couvrent chacune des différentes classes qui composent l'image. La sortie de la segmentation sémantique est, encore une fois, un masque de pixels dont la taille est égale à celle de l'image d'entrée.

2.2.2 Utilisation de vision par ordinateur

La Vision par ordinateur est omniprésente dans la technologie d'aujourd'hui. Il est donc compliqué de dresser une liste totalement exhaustive de ses domaines d'applications. En effet, de plus en plus, des domaines et produits basent leur construction et/ou leur développement technologique sur la vision par ordinateur [67].

La figure 2.1 ci-dessous présente les domaines d'application de la vision par ordinateur.

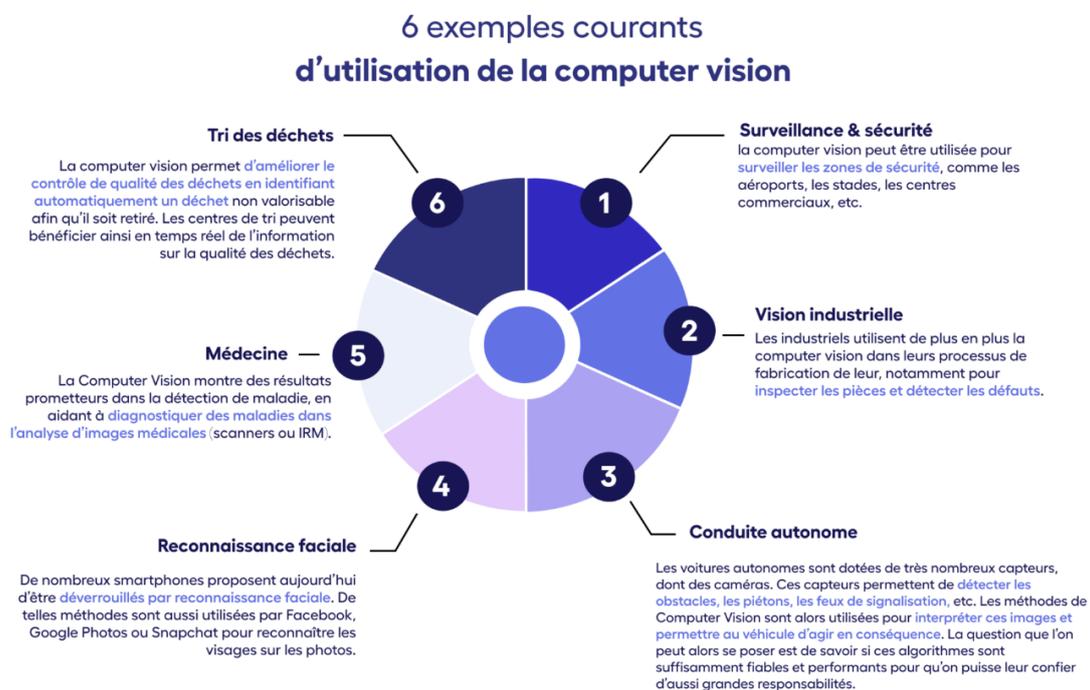


Figure 2.1 : Domaine d'application de la vision par ordinateur [67]

Chapitre 2 : Vision par ordinateur et segmentation d'instance

2.3 Détection d'objet

La détection d'objets [56] est considérée comme l'un des domaines remarquables de l'apprentissage profond et de la vision par ordinateur. La détection d'objets a été déterminée par de nombreuses applications en vision par ordinateur telles que le suivi d'objets, la récupération, la vidéosurveillance, le sous-titrage d'images, la segmentation d'images et un plus grand nombre d'autres applications. La figure 2.2 ci-dessous montre un exemple de détection d'objet.

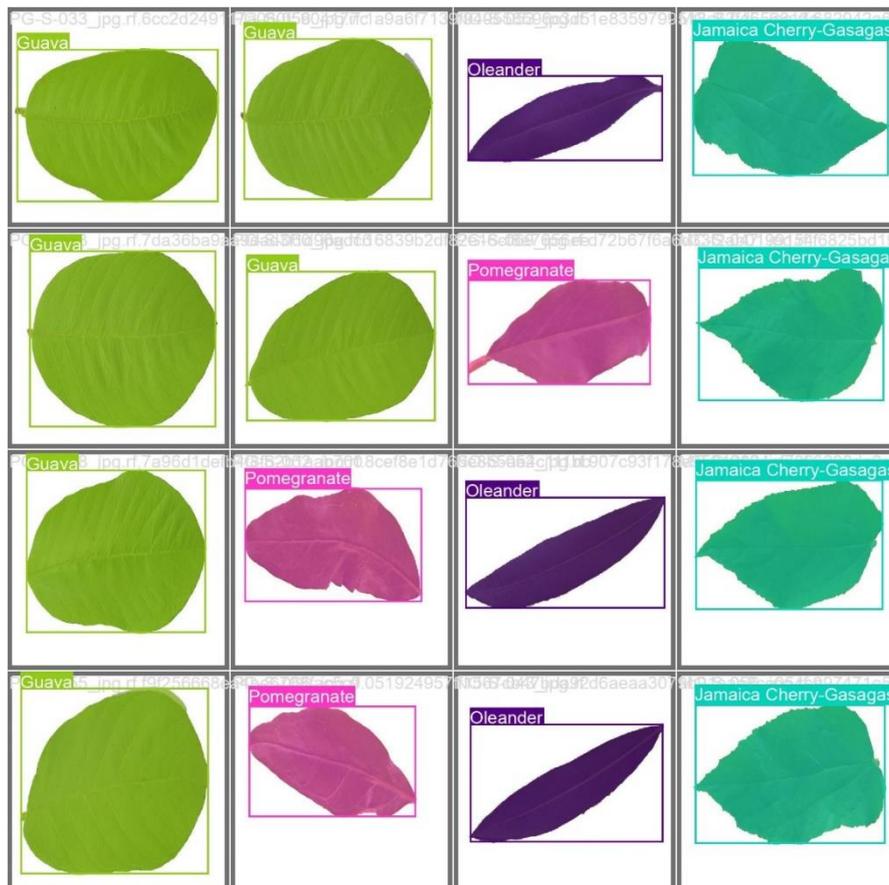


Figure 2.2 : Exemple de détection d'objet.

2.3.1 Les Techniques de détection d'objet

Le cadre de détection d'objets [53] est divisé en deux types de catégories. La figure 2.3 ci-dessous montre la catégorisation des diverses techniques de détection d'objets en deux groupes : les détecteurs traditionnels et les détecteurs basés sur l'apprentissage profond.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

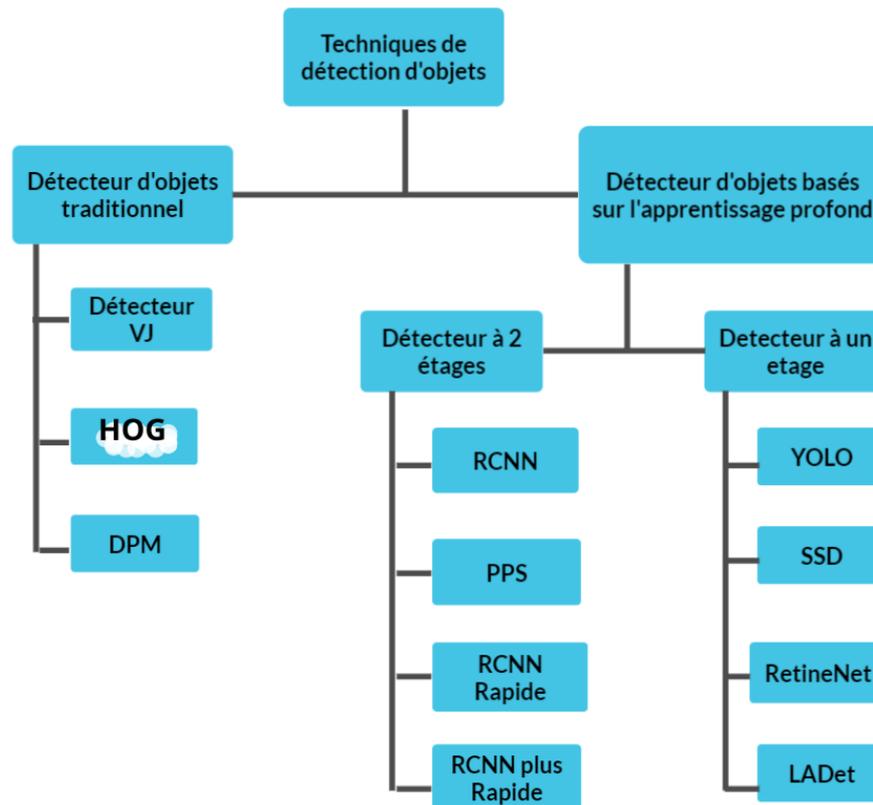


Figure 2.3 : Classification des techniques de détection d'objet.

➤ *Détecteurs d'objets traditionnels*

- *Viola Jones détecteur* : est un système de détection d'objets utilisant l'intelligence artificielle. Son objectif est de repérer des objets, en particulier des visages, dans les images. L'algorithme est à la fois performant et solide, ce qui le rend approprié pour les appareils à ressources restreintes.
- *HOG Detector* : Des histogrammes de gradients orientés sont calculés dans des cellules locales, capturant les orientations des bords de l'image. Ces histogrammes sont conçus manuellement pour capturer l'apparence des bords et des contours dans différentes orientations.
- *DPM Detector* : les modèles DPM prennent en compte les structures hiérarchiques basées sur les pièces, l'apparence et les modèles géométriques de chaque pièce sont définis manuellement. La capacité du modèle à prendre en compte les déformations et les relations spatiales est également conçue sur la base de connaissances préalables [68].

Chapitre 2 : Vision par ordinateur et segmentation d'instance

➤ *Détecteurs d'objets basés sur l'apprentissage profond*

❖ *Détecteurs à un étage*

- *SSD* : est un dispositif de détection d'objets à une étape qui utilise des réseaux de neurones convolutifs (CNN) afin de repérer des objets dans des images. Il adopte une méthode simple pour créer des boîtes englobantes et trier les objets en une seule fois, ce qui le rend rapide et performant.
- *RetinaNet* : est un modèle de détection d'objets en une étape qui utilise une fonction de perte focale pour remédier au déséquilibre de classe pendant l'entraînement. La perte focale applique un terme modulateur à la perte d'entropie croisée afin de concentrer l'apprentissage sur des exemples négatifs concrets.[73]
- *LADet* : est une technique de détection d'objets dans des images qui utilise une architecture de réseau de neurones convolutif (CNN). Il a été développé pour être léger et flexible afin de s'adapter aux applications de détection d'objets à grande vitesse et délicate.

❖ *Détecteurs à 2 étages*

- *SPP-Net*: est une architecture neuronale convolutionnelle qui utilise le regroupement pyramidal spatial pour supprimer la contrainte de taille fixe du réseau. Plus précisément, nous ajoutons une couche SPP au-dessus de la dernière couche convolutive. La couche SPP regroupe les fonctionnalités et génère des sorties de longueur fixe, qui sont ensuite introduites dans les couches entièrement connectées (ou d'autres classificateurs) [72].
- *Le R-CNN rapide, et le R-CNN plus rapide* : sont un système de détection d'objets qui utilise des réseaux de convolution neuronale (CNN) afin d'améliorer la vitesse et la précision de la détection d'objets dans des images.

2.4 Segmentation d'image

La segmentation d'images [39] est une tâche clé en vision par ordinateur et en traitement d'images avec des applications importantes telles que la compréhension de scènes, l'analyse d'images médicales, la robotique, vidéosurveillance, réalité augmentée et compression d'images, entre autres. Et de nombreux algorithmes de segmentation sont trouvés dans la littérature.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

La segmentation d'image est l'une des principales étapes du traitement d'image [48], dans laquelle toute image est subdivisée en plusieurs segments. Chaque segment représentera une sorte d'information à l'utilisateur sous forme de couleur, d'intensité ou de texture. Ainsi, il est important d'isoler les limites de toute image sous la forme de ses segments.

2.4.1 Les méthodes de segmentation d'image

Figure 2.4 présente quelques techniques ou méthodes de segmentation d'images :

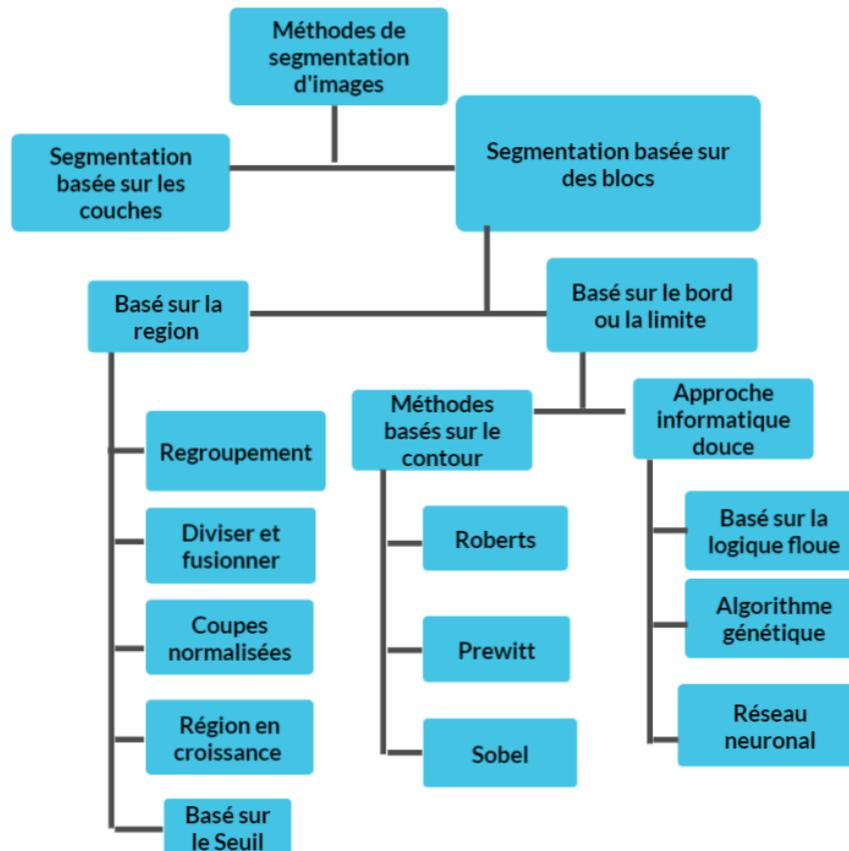


Figure 2.4 : Méthode de segmentation d'image [48].

2.4.2 Segmentation d'image en utilisant l'apprentissage profond

L'utilisation des techniques de l'apprentissage profond offre de nouvelles possibilités passionnantes pour la détection et la segmentation d'image. Ces approches permettent de tirer parti de l'explosion des données disponibles, comme les images satellites, les relevés de terrain et les bases de données d'images, afin de développer des outils puissants. Dans cette section nous présenterons quelques techniques pour la détection des plantes :

Chapitre 2 : Vision par ordinateur et segmentation d'instance

- CNN est une architecture d'algorithmes de l'apprentissage en profondeur contenant un ensemble de fonctions de transformation non linéaires. Ces réseaux sont une version avancée des réseaux de neurones artificiels (ANN) capables d'effectuer des opérations de traitement d'images telles que l'identification, la segmentation et la classification d'objets.
- **Processus typique de segmentation avec l'apprentissage profond** : un modèle basé sur un réseau neuronal convolutif (CNN) est appris qui compresse d'abord l'image source avec une pile de différentes couches de convolution, d'activation et de pooling. La représentation latente compressée est développée par les opérations inverses. Le réseau reste entraînable de bout en bout. Au moment du test, une passe avant fournit les étiquettes de segmentation [17].

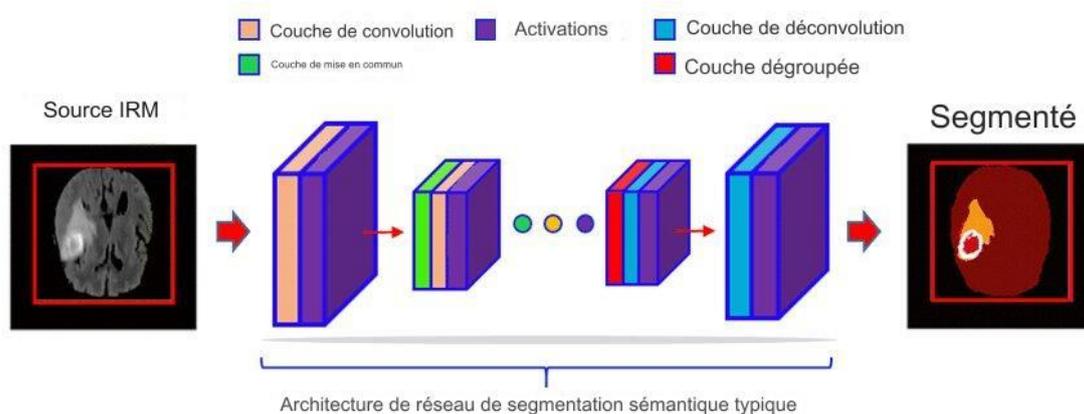


Figure 2.5 : CNN Architecture pour la Segmentation sémantique [17].

- **Modèles basés sur un codeur-décodeur** : Fondamentalement, les architectures de style encodeur et décodeur utilisent les mêmes couches d'auto-attention pour coder les jetons de mots. Cependant, la principale différence réside dans le fait que les encodeurs sont conçus pour apprendre des intégrations pouvant être utilisées pour diverses tâches de modélisation prédictive telles que la classification. En revanche, les décodeurs sont conçus pour générer de nouveaux textes, par exemple en répondant aux requêtes des utilisateurs [55]. Les modèles basés sur un codeur-décodeur peuvent être regroupés en deux catégories :
 - **A. Modèles d'encodeur-décodeur pour la segmentation générale** : Il se compose de deux parties, un encodeur et un décodeur.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

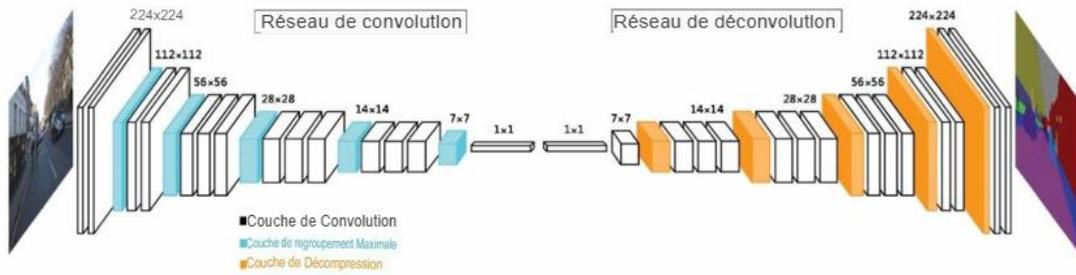


Figure 2.6 : Segmentation sémantique déconvolutive suivie d'un réseau convolutionnel de couche VGG-16 [46].

- **B. Modèles d'encodeur-décodeur pour la segmentation d'images médicales et biomédicales :** U-Net et V-Net sont les deux architectures les plus populaires utilisées dans la segmentation d'images médicales/biomédicales.
 - **U-Net :** Initialement développé pour la segmentation d'images de cellules biologiques, U-Net est un modèle de segmentation sémantique extrêmement efficace. Grâce à sa structure en U, composée d'une partie de contraction et d'une partie d'expansion, elle peut prendre en compte à la fois des informations à grande et à petite échelle.
 - **V-Net :** est une variante de U-Net qui permet de traiter des images en 3D, telles que les scans médicaux. Outre les connexions skip horizontales, V-Net intègre également des connexions skip verticales afin de garantir une meilleure conservation des informations 3D tout au long de son réseau.

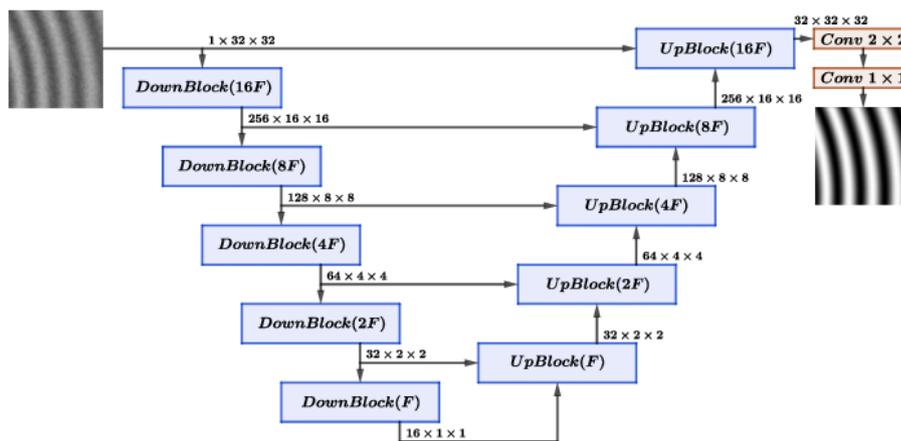


Figure 2.7: U-Net et V-Net [14].

Les architectures U-net et V-net se ressemblent au niveau des blocs. Cependant, le nombre de filtres par bloc est inversement réparti : U-net est destiné à la segmentation

Chapitre 2 : Vision par ordinateur et segmentation d'instance

(classification) d'images tandis que V-net est conçu pour la reconstruction d'images (régression).

- **Modèles basés sur des réseaux multi-échelles et pyramidaux :** Feature Pyramid Network (FPN) est le modèle le plus populaire dans cette catégorie. Initialement, il a été développé pour la détection d'objets, mais a ensuite été utilisé également pour la segmentation d'images. Le Feature Pyramid Network (FPN) est un système d'extraction de fonctionnalités spécialement conçu pour faire face à un tel concept de pyramide avec précision et rapidité. Il est utilisé en remplacement de l'extracteur de caractéristiques des détecteurs tels que R-CNN plus rapide et produit plusieurs couches de cartes de caractéristiques (cartes de caractéristiques multi-échelles) qui fournissent des informations de meilleure qualité que la pyramide de caractéristiques traditionnelle pour détecter objets.

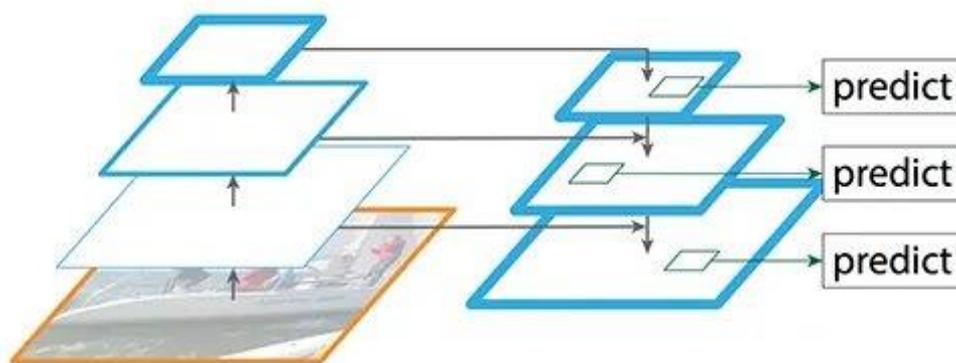


Figure 2.8: Réseaux multi-échelles et pyramidaux [45].

- **Modèles basés sur R-CNN (par exemple segmentation) :** Le réseau convolutif régional (RCNN) est un modèle très populaire abordant le problème de la segmentation des instances. Il effectue simultanément les tâches de détection d'objets et de segmentation sémantique. Le modèle de détection d'objets R-CNN utilise des CNN de grande capacité pour ascendant les propositions de régions pour localiser et segmenter les objets. Il emploie une approche sélective afin de repérer un certain nombre de zones d'objets englobantes potentielles (« zones d'intérêt »), puis extrait séparément les caractéristiques de chaque zone à des fins de classification.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

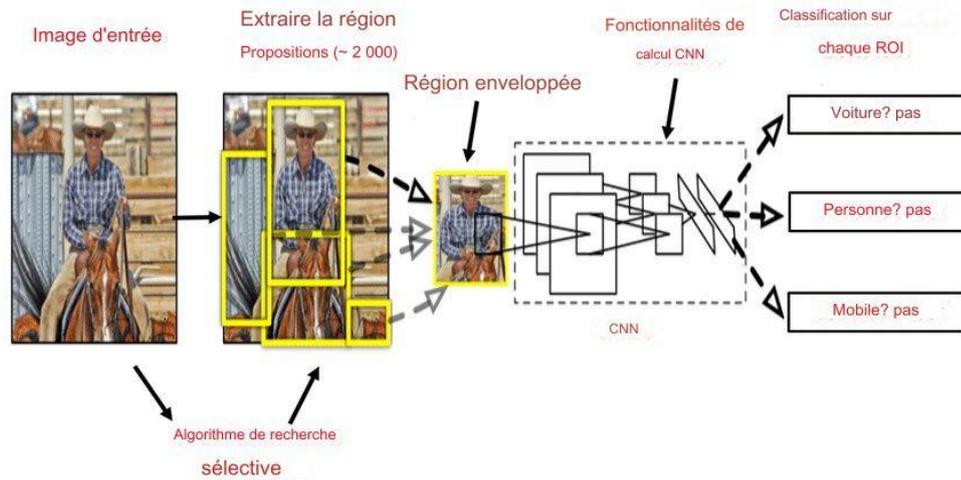


Figure 2.9: R-CNN [20].

2.4.3 Types de tâches de segmentation d'images

Voici les types de tâches de segmentation d'images les plus courants [49] :

- **Segmentation sémantique** : vise à prédire une étiquette de classe par pixel sans faire de distinction entre plusieurs instances appartenant à la même catégorie.
- **Segmentation d'instance** : issue de la détection d'objets, la segmentation d'instance prédit les probabilités de classe, les cadres de délimitation et les masques d'instance dans une image. Comme son nom l'indique, la segmentation des instances identifie des instances distinctes de la même catégorie.
- **Segmentation panoptique** : la segmentation panoptique unifie la segmentation sémantique et instance. Il est capable d'attribuer une étiquette de classe pour chaque pixel d'une image et un ID d'instance pour tous les pixels appartenant à un objet.



Figure 2.10: Les types de segmentation d'images [71].

Chapitre 2 : Vision par ordinateur et segmentation d'instance

2.4.4 Avantages et inconvénients pour chaque type de segmentation d'image

La segmentation sémantique, par instance et panoptique sont les trois types de segmentation d'image. Voici les avantages et inconvénients de chaque type :

| Type | Avantage | Inconvénient |
|-------------------------|--|--|
| Segmentation d'instance | <ul style="list-style-type: none">• Détection d'objets plus précise et localisation [16].• Précision améliorée dans la reconnaissance et la classification des objets [32].• Meilleure performance dans des tâches telles que : suivi d'objet et segmentation d'image [32].• Utilisation plus efficace des ressources informatiques [32]. | <ul style="list-style-type: none">• Coût de calcul et besoins en mémoire élevés, qui peuvent rendre difficile son utilisation dans des applications en temps réel [62].• Difficulté à gérer l'occlusion et le chevauchement des objets, ce qui peut entraîner des erreurs dans la détection et la segmentation des objets [33].• Complexité de la formation et de la mise au point des modèles, ce qui peut nécessiter de grandes quantités de données étiquetées et ressources informatiques [1]. |

Table 2.1 – Avantage et Inconvénient de la segmentation d'instance.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

| Type | Avantage | Inconvénient |
|-------------------------|---|--|
| Segmentation sémantique | <ul style="list-style-type: none"> • Capacité à fournir une compréhension détaillée du contenu d'une image, ce qui est utile dans des domaines tels que l'imagerie médicale et la conduite autonome [54]. • De meilleures performances dans des tâches telles que la reconnaissance et la classification d'objets, ce qui est important dans des domaines tels que l'agriculture et la surveillance environnementale [27]. • Capacité à générer des données plus détaillées et précises pour les modèles d'apprentissage, ce qui est important dans des domaines tels que la santé et la finance [28]. | <ul style="list-style-type: none"> • Coût de calcul et besoins en mémoire élevés [19]. • Difficulté à atteindre une grande précision dans les applications en temps réel [41]. • Difficulté à gérer de grands ensembles de données [7]. |

Table 2.2 – Avantage et Inconvénient de la Segmentation Sémantique.

Chapitre 2 : Vision par ordinateur et segmentation d'instance

| Type | Avantage | Inconvénient |
|-------------------------|--|---|
| Segmentation Panoptique | <ul style="list-style-type: none">• Permet de mesurer la cohérence des annotations humaines [5].• Aucun post-traitement de la sortie panoptique n'est nécessaire [8].• Combine la segmentation sémantique et la segmentation d'instance [5]. | <ul style="list-style-type: none">• Il peut être affecté par un déséquilibre de classes, où certaines classes d'objets sont sous-représentés dans les données sur la formation [30].• Il peut avoir du mal à détecter des objets partiellement obstrués ou ayant des formes complexes [30].• Il peut être affecté par des biais dans les données de formation ou l'architecture du modèle [29]. |

Table 2.3 – Avantage et Inconvénient de la Segmentation Panoptique.

2.4.5 Pourquoi La segmentation d'instance ?

La catégorisation d'images est un domaine captivant de l'intelligence artificielle qui permet de saisir et de repérer le contenu visuel des images. La technologie offre de multiples opportunités d'applications novatrices dans des secteurs tels que la robotique, la médecine et la sécurité.

La segmentation des instances est devenue l'un des domaines relativement importants, complexes et difficiles de la recherche en vision industrielle. Destiné à prédire l'étiquette de classe d'objet et le masque d'instance d'objet spécifique au pixel, il localise différentes classes d'instances d'objet présentes dans diverses images. La segmentation des instances vise à aider en grande partie la robotique, la conduite autonome, la surveillance, etc..[1].

Chapitre 2 : Vision par ordinateur et segmentation d'instance

2.5 Conclusion

Ce chapitre était centré sur le sujet de la vision par ordinateur, avec un accent particulier sur la technique de reconnaissance, et la segmentation des instances comme domaines clés. Dans le troisième chapitre, on discutera de son utilisation dans notre application.

Chapitre 3 : Conception & Réalisation

Chapitre 3 : Conception & Réalisation

3.1 Introduction

Ce chapitre présentera l'architecture proposée pour développer le modèle et fournira un aperçu complet de l'ensemble de données utilisé. Le modèle pour identifier les feuilles des plantes a été développé en mettant en œuvre des techniques d'apprentissage profond.

3.2 Architecture du Système

La première étape de l'architecture du système sera la collecte d'images où les images sont collectées et stockées pour l'utilisation du système. L'étape suivante est le prétraitement de l'image où les images sont annotées. La prochaine étape serait l'augmentation des données où les images sont augmentées à l'aide de la méthode d'augmentation des données tel que la Rotation, Flip. Après cela, nous allons définir notre modèle sélectionné et Ensuite, entraînez le modèle en utilisant des données d'entraînement adéquates. Utilisez des jeux de données de validation et de test pour évaluer le modèle. Évaluez la performance du modèle en utilisant des indicateurs tels que la précision, le rappel et le F1-score. Ensuite, nous allons l'intégrer dans notre application. La figure 3.1 présente l'architecture du système.

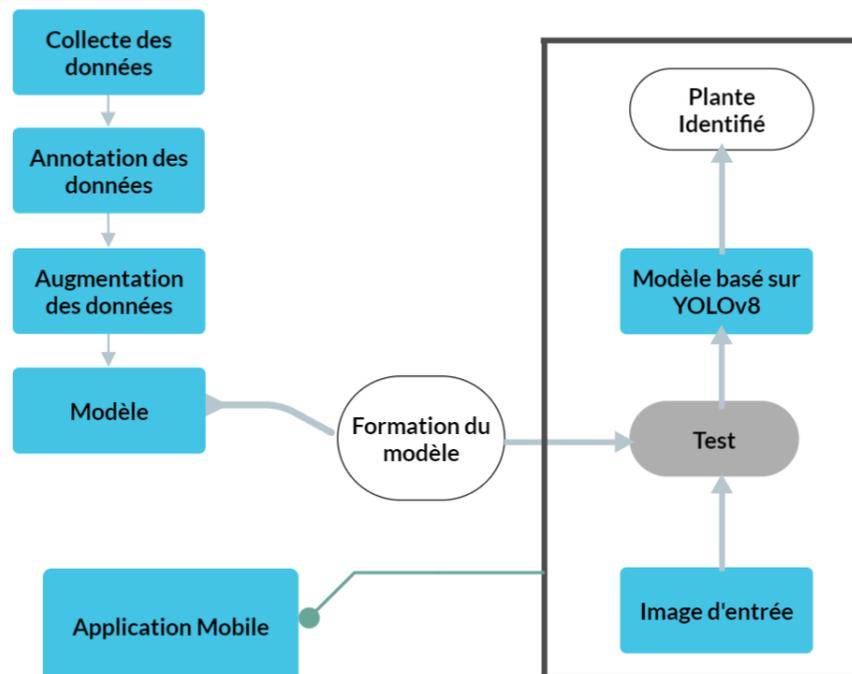


Figure 3.1 : Architecture du système.

Chapitre 3 : Conception & Réalisation

3.3 Collecte et prétraitement de données

Dans toute approche basée sur les données, la collecte de données et le prétraitement sont des étapes critiques qui nécessitent une attention particulière aux détails. La qualité des données collectées et la façon dont elles sont traitées peuvent avoir un impact significatif sur la précision et l'efficacité des modèles construits. Dans les opérations de la chaîne d'approvisionnement, la collecte de données et le prétraitement sont particulièrement difficiles en raison des grandes quantités de données générées à partir de diverses sources, notamment des fournisseurs, des fabricants, des distributeurs et des détaillants.

Ces données peuvent être structurées ou non structurées, ce qui les rend encore plus difficiles à intégrer et à analyser.

Par conséquent, il est essentiel de collecter et de prétraiter les données d'une manière qui garantit sa précision, son exhaustivité et sa cohérence [59].

| N | Nom de la plante | Exemple d'images |
|---|------------------|---|
| 1 | Betel |  |
| 2 | Neem |  |
| 3 | Tulsi |  |
| 4 | Basale |  |
| 5 | Guava |  |

Table 3.1 : Images de quelques plantes.

Chapitre 3 : Conception & Réalisation

Les images de ce projet ont été recueillies sur Google Images, où les plantes de chaque image ont été sélectionnées, ainsi que d'autres images sur Kaggle (Medicinal leaf dataset).

Constituer de 2756 images.

3.3.1 Annotation des données

Afin de préparer les images collectées pour la formation du modèle de segmentation d'instance, la plate-forme Roboflow ⁶ a été utilisée pour l'annotation manuelle des images. Cet outil utilise l'apprentissage automatique lui-même, offrant des fonctionnalités telles que la gestion des données d'image, d'annoter et d'étiqueter des ensembles de données qui peut être assister par une IA, d'appliquer des prétraitements et des augmentations, de convertir les formats des fichiers d'annotation, de créer un modèle de vision par ordinateur en un seul clic et de déployer des modèles via une API ou en périphérie. Plusieurs tâches de Computer Vision sont prises en charge par Roboflow : (Classification, Détection d'objets, Segmentation d'instance).

Les images de l'ensemble de données ont été annotées avec des étiquettes identifiant l'emplacement spécifique des feuilles des plantes.

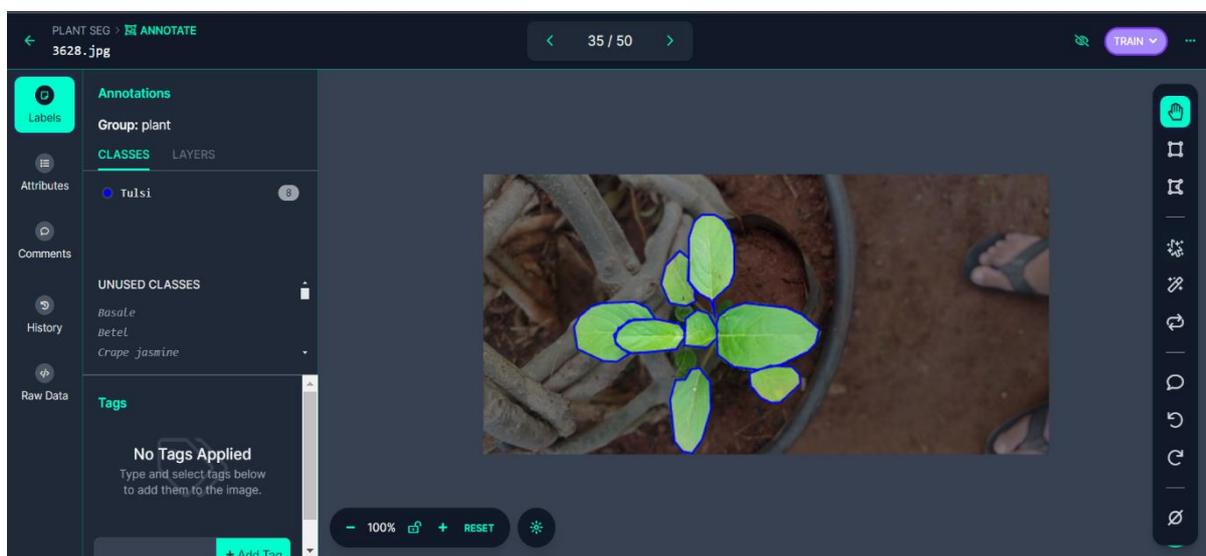


Figure 3.2 : Annotation avec outils Roboflow

3.3.2 Augmentation des données

Consiste à augmenter le nombre d'images pour obtenir un résultat plus précis. C'est une technique pour créer artificiellement un nouvel ensemble de données à partir de données existantes. L'augmentation des données d'image comprend diverse série opérations pour

⁶ <https://roboflow.com/>

Chapitre 3 : Conception & Réalisation

manipuler l'image telles que les décalages, les zooms, les retournements sont les méthodes utiliser dans ce projet, cette approche invariante et aide également le modèle à apprendre des caractéristiques invariantes.

La Figure 3.3 ci-dessous présente un Echantillon d'augmentation.

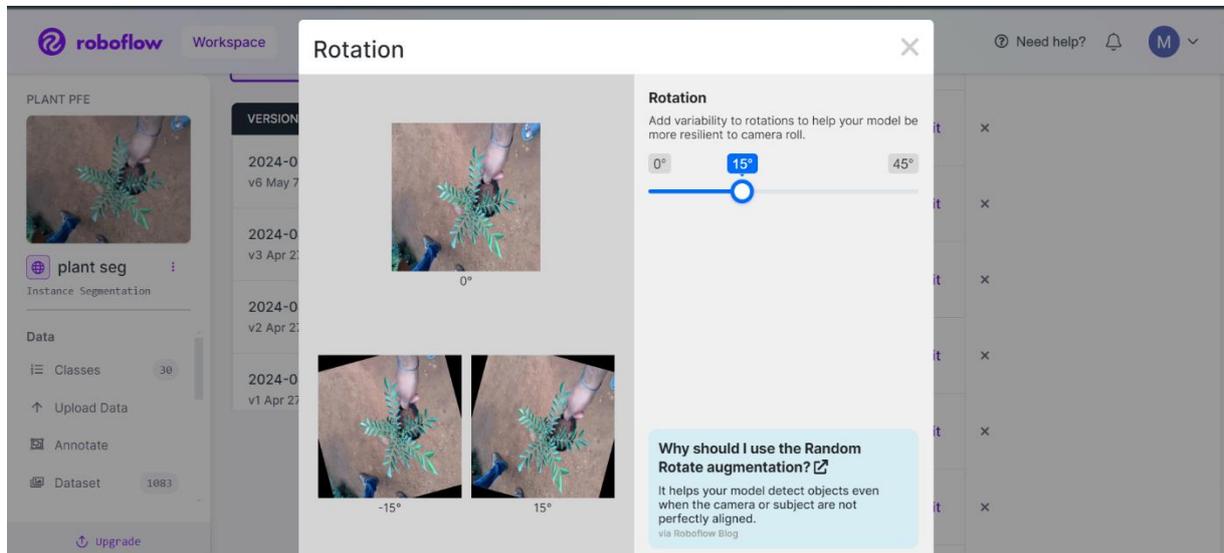


Figure 3.3 : Echantillon d'augmentation

3.4 YOLO

You Only Look Once, ou YOLO est un algorithme de détection d'objets en temps réel populaire. YOLO combine ce qui était autrefois un processus en plusieurs étapes, utilisant un seul réseau neuronal pour effectuer à la fois la classification et la prédiction des cadres de délimitation des objets détectés [9].

L'un des principaux avantages de YOLO est sa vitesse d'inférence rapide, qui lui permet de traiter les images en temps réel. Il convient parfaitement aux applications telles que la vidéosurveillance, les voitures autonomes et la réalité augmentée. De plus, YOLO a une architecture simple et nécessite un minimum de données de formation, ce qui facilite sa mise en œuvre et son adaptation à de nouvelles tâches [76].

Chapitre 3 : Conception & Réalisation

3.4.1 YOLO Architecture

L'architecture du modèle CNN qui constitue l'épine dorsale de YOLO [76].

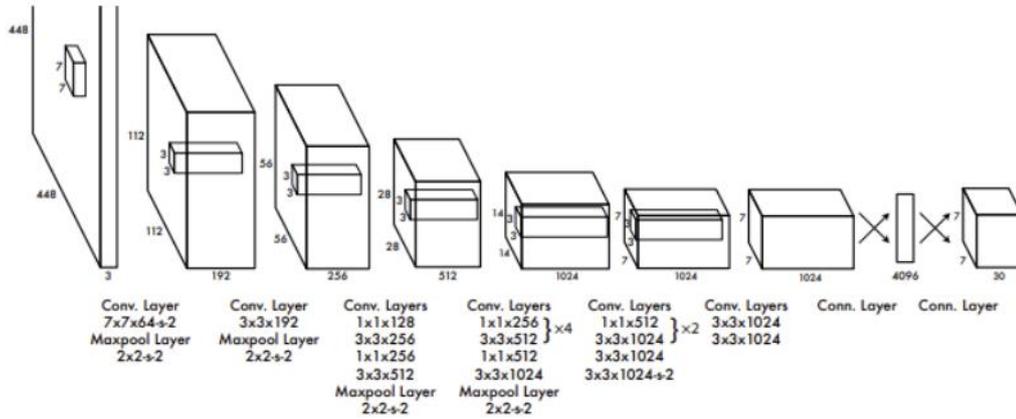


Figure 3.4 : YOLO Architecture [76].

3.4.2 YOLO Versions

YOLO a subi des évolutions à travers différentes versions : YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, YOLOv7, YOLOv8 et YOLOv9. Chaque version améliorée de la précédente en termes de vitesse, de précision et de fonctionnalité.

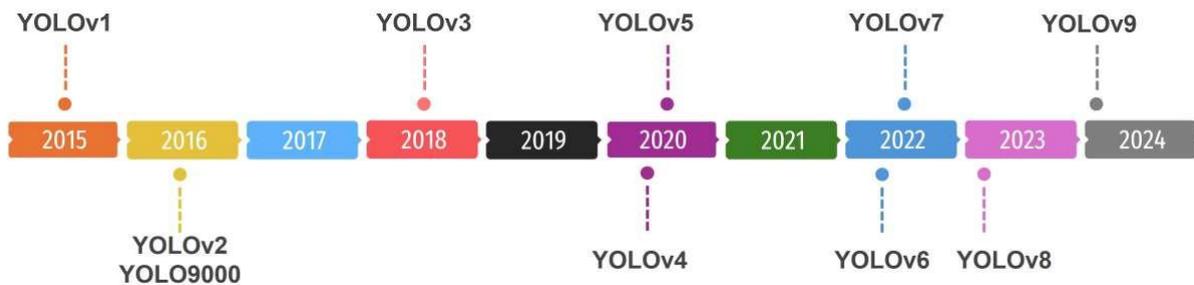


Figure 3.5 : YOLO évolution [64].

3.4.3 YOLOv8

L'un des algorithmes de détection d'objets les plus populaires est You Only Look Once (YOLO), qui a connu plusieurs itérations depuis son introduction en 2016. YOLOv8 est très rapide et très précis, obtenant des résultats de pointe sur plusieurs ensembles de données de référence [4].

Chapitre 3 : Conception & Réalisation

Notre étude actuelle utilise le modèle YOLOv8, qui s'est révélé efficace pour identifier les plantes dans notre ensemble de données.

3.4.4 Pourquoi devons-nous utiliser YOLOv8 ?

YOLOv8 est un modèle très puissant de la famille très influente de modèles utilisés pour détection d'objets et segmentation d'images. Il présente une nouvelle architecture, de nouvelles couches convolutives et une nouvelle tête de détection.

Il est également nettement plus rapide et plus précis que les versions précédentes de YOLO, ce qui en fait un excellent choix pour la détection d'objets. De plus, YOLOv8 prend en charge les derniers algorithmes de vision par ordinateur, y compris la segmentation des instances, ce qui permet la détection de plusieurs objets dans une image [34].

3.5 YOLO pour La segmentation d'instance

Au cours des dernières années, la communauté de la vision a fait de grands progrès, par exemple la segmentation d'instances, en partie en s'appuyant sur des parallèles puissants avec le domaine bien établi de la détection d'objets. Approches de pointe en matière de segmentation d'instance, telles que Mask R-CNN s'appuient directement sur les avancées en matière de détection d'objets telles que R-CNN plus rapide. Pourtant, ces méthodes se concentrent principalement sur les performances plutôt que sur la vitesse, laissant la scène dépourvue de parallèles de segmentation d'instance avec les détecteurs d'objets en temps réel comme SSD et YOLO.

Cependant, la segmentation des instances est difficile, bien plus difficile que la détection d'objets. Les détecteurs d'objets à un étage comme SSD et YOLO sont capables d'accélérer les détecteurs à deux étages existants comme RCNN plus rapide en supprimant simplement le deuxième étage et en compensant la performance perdue par d'autres moyens [13].

3.6 Segmentation YOLOv8 pour la détection de plantes médicinales

Le modèle YOLOv8-Seg (ultralytics-8.0.57) se compose de deux composants principaux : le squelette et la tête (qui peuvent être divisés en cou et segment). Le GitHub propose cinq modèles réduits différents du réseau, à savoir YOLOv8n-Seg, YOLOv8s-Seg, YOLOv8m-Seg, YOLOv8l-Seg et YOLOv8x-Seg [51].

Chapitre 3 : Conception & Réalisation

Dans ce projet :

- Des expériences ont été réalisées sur le modèle YOLOv8-Seg.
- Les données étant importées dans Roboflow, on doit exporter la base dans notre espace de travail Google Collab.
- Une fois que les données d'entraînement sont prêtes et que le modèle est configuré,

Vous pouvez lancer le processus d'entraînement. Cela implique de définir les paramètres d'entraînement, tels que le nombre d'époques, le taux d'apprentissage et la taille des lots, et de surveiller attentivement les métriques de performance tout au long du processus.

- Puis une fois le modèle entraîné et que le résultat de l'entraînement est satisfaisant alors il faudra valider le modèle.

- Et le tester sur la base de test.

- La précision, le rappel, la matrice de confusion seront des mesures qui nous permettront d'évaluer ce modèle (on les détaillant dans le chapitre suivant).

3.7 Intégration de la segmentation YOLOv8 dans l'application mobile

Une fois le modèle entraîné et évalué avec succès, il doit être intégré dans un système plus large, tel qu'une application médicale ou une plateforme d'identification de plantes. Cette étape implique de s'assurer que le modèle est facilement accessible et qu'il peut être utilisé de manière transparente par les utilisateurs finaux.

Exigences évoluent, le modèle doit être régulièrement mis à jour et amélioré. Cela peut nécessiter de nouvelles phases d'entraînement, de test et de déploiement pour garantir des performances optimales dans le temps.

Chapitre 3 : Conception & Réalisation

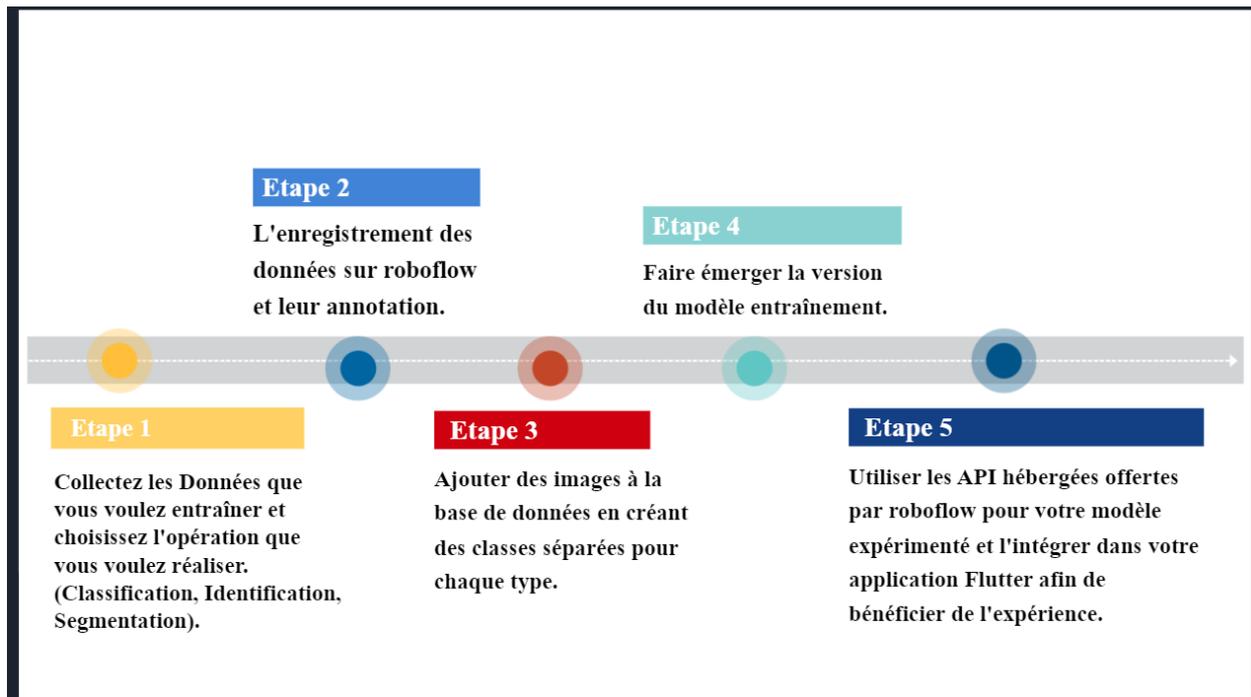


Figure 3.6: Les Etapes de déploiement sur votre Application

Le modèle est téléchargé à l'aide de l'API d'inférence de Roboflow et utilisé pour l'inférence. Les modèles de vision par ordinateur, tels que la détection d'objets, la classification, la segmentation d'instances et la segmentation sémantique, peuvent être déployés grâce à l'API d'inférence. Il est possible d'exécuter directement les modèles via l'API, sans avoir besoin de les télécharger localement.

L'API est spécialement développée pour être intégrée de manière aisément à des applications, offrant ainsi la possibilité de repérer des objets, de classer des images et de segmenter des instances sans avoir à gérer les détails de l'inférence.

Avec un GPU NVIDIA, il est possible d'accélérer votre inférence en installant les dépendances CUDA compatibles.

3.8 Conclusion

Ce chapitre était le plus essentiel, car il expliquait les principes fondamentaux, le modèle utilisé, architecture du système, comment on a pu déployer le modèle sur roboflow. La discussion sur les résultats trouvés sera l'objectif principal du prochain chapitre.

Chapitre 4 : Implémentation & Résultats

Implémentation & Résultats

4.1 Introduction

Ce chapitre présente les informations essentielles concernant les logiciels et le matériel employés dans le cadre de ce projet. Il fournira quelques informations sur les résultats obtenus ainsi qu'une comparaison avec un modèle différent.

4.2 Conception Logiciel

Python est un langage de programmation interprété, de haut niveau et à usage général. Créée par Guido van Rossum et publiée pour la première fois en 1991, la philosophie de conception de Python met l'accent sur la lisibilité du code avec son utilisation notable d'espaces blancs importants. Ses constructions de langage et sa méthodologie de localisation des éléments visent à aider les ingénieurs logiciels à composer du code légitime pour des projets de petite et grande envergure.

Python est un langage de programmation de haut niveau qui a attiré l'attention des programmeurs grâce à sa philosophie de conception et à sa fiabilité des codes et de la syntaxe du langage de programmation Python qui nous permet d'exprimer des concepts par rapport à tout autre langage de programmation [10].

Bibliothèques Python utilisées dans ce projet :

Ultralytics YOLOv8 est développé par Ultralytics, est un outil de pointe (SOTA) qui s'appuie sur le succès des versions précédentes de YOLO et introduit de nouvelles fonctionnalités et améliorations pour améliorer encore les performances et la flexibilité. YOLOv8 est conçu pour être rapide, précis et facile à utiliser, ce qui en fait un excellent choix pour un large éventail de détection d'objets, segmentation d'images et classification d'images tâches [58].

Autres outils :

GoogleCollab offre la possibilité d'écrire et d'exécuter du code Python dans un navigateur web sans avoir besoin d'installer Python ou d'autres bibliothèques sur votre ordinateur. Il permet d'accéder gratuitement à des processeurs graphiques (GPU) et des unités de traitement de tenseur, ce qui est très bénéfique pour les tâches qui nécessitent beaucoup de calcul, telles que l'utilisation de modèles IA avancés.

Flutter un SDK open source créé par Google, offre la possibilité de concevoir des expériences utilisateur agréables et rapides pour les applications mobiles, Web et de bureau en utilisant une

Implémentation & Résultats

seule base de code. Son objectif est d'être productif, flexible et ouvert, offrant aux développeurs la possibilité de concevoir aisément des applications belles et rapides.

Dart est un langage de programmation open source à usage général développé par Google. Il prend en charge le développement d'applications côté client et côté serveur. Cependant, il est largement utilisé pour le développement d'applications Android, d'applications iOS, d'IoT (Internet des objets) et d'applications Web à l'aide du Flutter Framework [60].

Android Studio est un outil de développement intégré (IDE) spécialement conçu pour la création d'applications mobiles Android.

Roboflow offre des outils de vision par ordinateur permettant de gérer des ensembles de données, d'entraîner des modèles et de les mettre en œuvre. Pour intégrer ses fonctionnalités dans des pipelines de vision par ordinateur, elle propose des API et des SDK.

Firebase est une plateforme de développement d'applications web et mobiles soutenue par Google, pour aider les développeurs à offrir des expériences d'applications plus riches. On parle de Backend As A Service (BaaS). Firebase gère sa propre infrastructure avec un bel ensemble d'outils pour simplifier le travail du développeur en lui fournissant des kits de développement et un tableau de bord en ligne [57].

4.3 Matériel

Voici la description de l'ordinateur qui a été utilisé pour construire le modèle :

- PC : HP Probook 445R G6
- Système d'exploitation : Windows 11 Professionnel
- PROCESSEUR : AMD Ryzen 5 3500U 2.10GHz
- MÉMOIRE : 8Go
- GRAPHIQUE : Radeon Vega Mobile Gfx.

Implémentation & Résultats

4.4 Interface utilisateur

La première étape du flux du système consiste à télécharger une image qui sera reconnue par le système. L'image chargée est ensuite identifiée en fonction du modèle. Les détails sont les sorties du identifier.

Nous avons choisi le nom « Medicinal plant in algeria » car il n'avait jamais été utilisé auparavant, et pour indiquer dans le contexte qu'il s'agit d'application d'identification d'image de la flore algérienne. Pour la conception, nous avons utilisé Figma, qui est un éditeur de graphiques vectoriels et un outil de prototypage. Il est principalement basé sur le Web, mais nous l'avons utilisé pour la conception de l'application et dispose de fonctionnalités hors ligne. J'ai choisi un design minimaliste et orienté végétal.

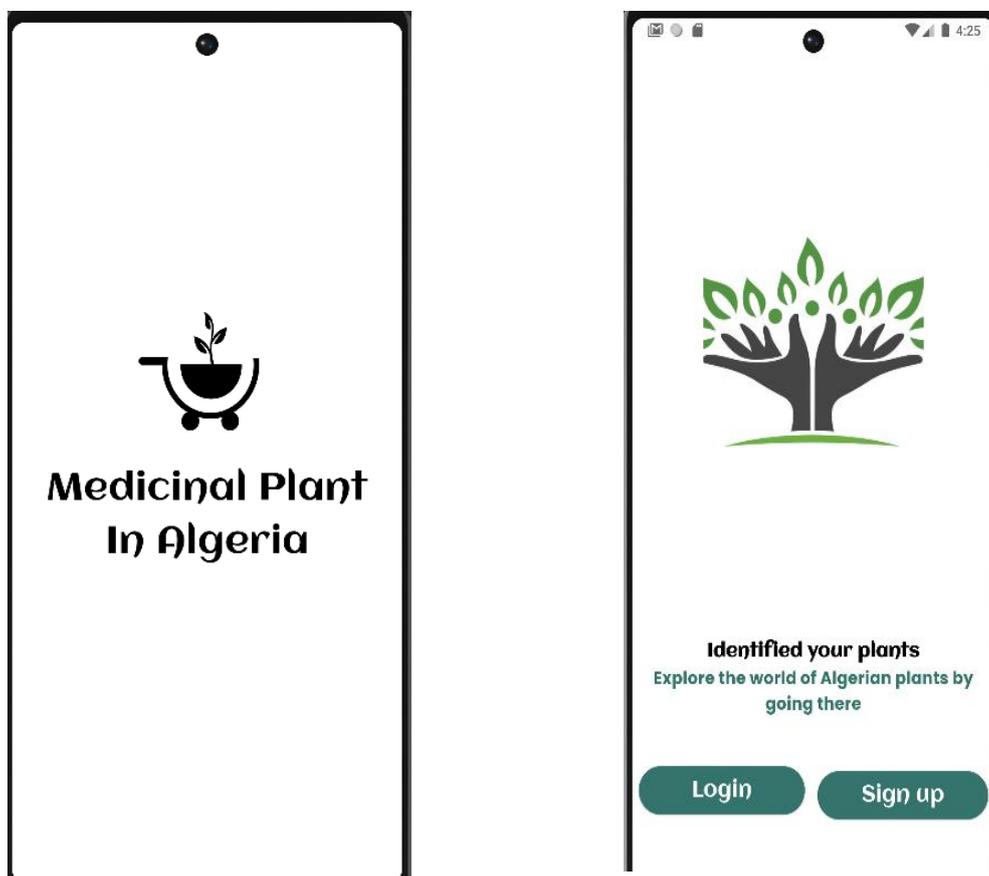


Figure 4.1 : Vitrine des pages d'intégration

4.5 Comment fonctionne l'identification des images dans l'application

Les graphiques suivants fournissent un résumé rapide de l'identification des images, affirmant simplement que tout est basé sur des icônes :

Implémentation & Résultats

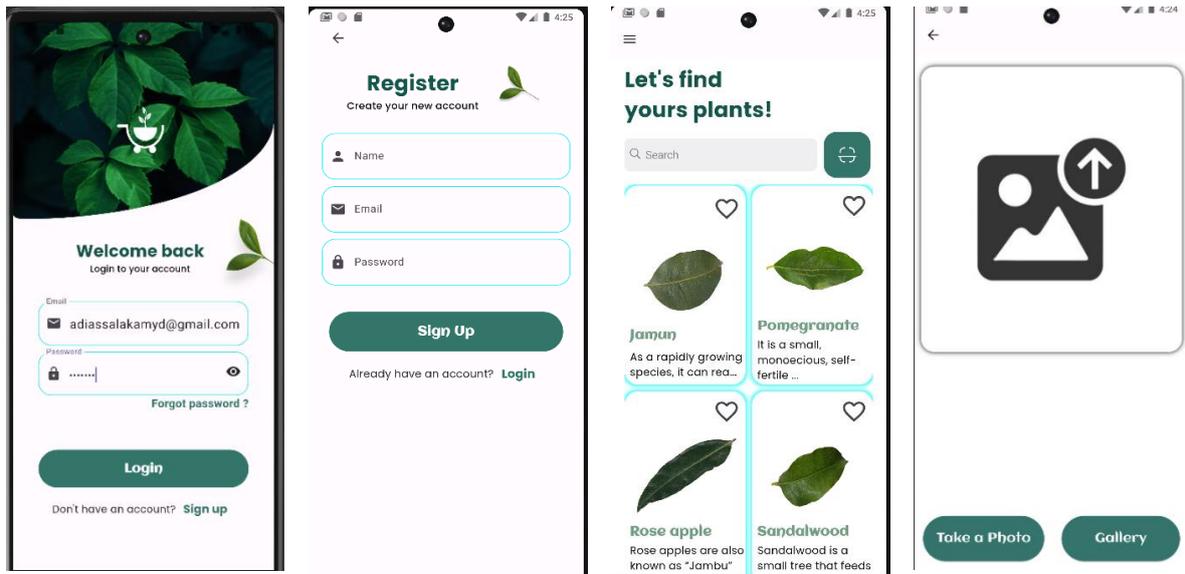


Figure 4.2 : Design des pages principales

4.6 Heuristique utilisée dans la conception

Les heuristiques sont un ensemble de principes fondamentaux utilisés dans les interactions interactives. Conception pour garantir la convivialité et offrir un moyen rapide et facile de découvrir les problèmes. Les problèmes et les solutions potentielles de la conception [75].

J'ai pu inclure certaines heuristiques connues dans la conception, notamment

- **Cohérence** : le design est resté cohérent sur l'ensemble de l'application. et les actions étaient dans un ordre logique que vous pouvez voir dans les photos Ci-dessous
- **Reconnaissance plutôt que rappel** : au lieu de mots et complexes structures, la conception reposait sur des images et des icônes, ce qui facilitait la tâche. L'utilisateur doit identifier plutôt que rappeler, ce qui entraîne une mémoire utilisateur inférieure pression.
- **Design esthétique et minimaliste** : les interfaces sont simples et offrent peu d'informations, permettant à l'utilisateur de se concentrer uniquement sur l'essentiel.

Implémentation & Résultats



Figure 4.3 : Tutoriel de l'application

4.7 Création de notre modèle d'apprentissage profond

Dans un premier temps, il est impératif d'installer la bibliothèque Ultralytics dans le but de construire et d'instruire un modèle de segmentation d'instance YOLOv8.

Ensuite, l'ensemble de données est importé dans le répertoire de travail actuel pour permettre l'accessibilité par le modèle à des fins de formation.

4.7.1 Affiner le modèle YOLOv8 pré-entraîné

Étant donné que YOLOv8 est un modèle pré-entraîné, il est impératif de l'affiner pour optimiser ses performances par rapport à notre ensemble de données spécifique. L'ajustement de ses hyperparamètres, notamment le taux d'apprentissage, la taille du lot et la régularisation, est nécessaire pour optimiser les performances du modèle.

Implémentation & Résultats

En apprentissage automatique, le terme « époques » fait référence au nombre d'itérations qu'un modèle subit pour apprendre de chaque image. Le terme « lot » fait référence au nombre d'images chargées et traitées lors de chaque itération du processus d'apprentissage. De plus, « imgsz » indique la taille de l'image utilisée dans le modèle.

4.7.2 Entraîner le modèle

La formation du modèle peut commencer après avoir apporté les ajustements nécessaires à l'ensemble de données. Google collab a été utilisée en raison de la nécessité d'un GPU pour accélérer le processus de formation. Au cours du processus de formation, le modèle améliore ses capacités de détection et de segmentation d'objets, améliorant progressivement sa précision.

```
epochs=30
workers=1
batch=8
imgsz=640
```

Le fichier nommé data.yaml sert de représentation des informations de l'ensemble de données, englobant les chemins d'images pour la formation, la validation, et les tests ainsi que les noms des classes.

```
train: ../train/images
val: ../valid/images
test: ../test/images

nc: 10
names: ['Basale', 'Betel', 'Fenugreek', 'Guava', 'Karanda', 'Mango', 'Neem', 'Tulsi', 'jasmine', 'lemon']

roboflow:
  workspace: plant-pfe
  project: plant-seg-fbe0f
  version: 28
  license: CC BY 4.0
  url: https://universe.roboflow.com/plant-pfe/plant-seg-fbe0f/dataset/28
```

Figure 4.4 : data.yaml

En se basant sur les données de 1184 images après leur augmentation. La figure 4.5 illustre la division des données d'entraînement, de validation et de test.

Implémentation & Résultats

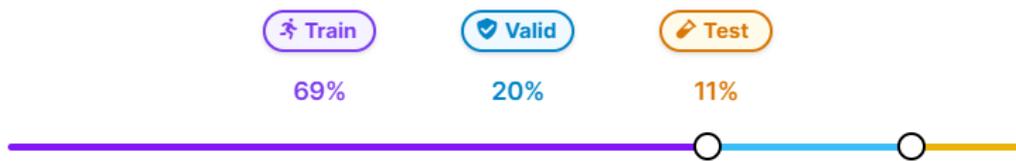


Figure 4.5 : Pourcentages des données d'entraînement, de validation et de test

4.8 Résultat et discussion

Dans cette étude, nous allons exposer les résultats obtenus et les interpréter, tout en les comparant avec ceux d'autres chercheurs dans le domaine.

4.8.1 La moyenne de la précision moyenne

La moyenne de la précision moyenne (mAP) est une métrique utilisée pour évaluer les modèles de détection d'objets tels que R-CNN rapide, YOLO, Mask R-CNN, etc. La moyenne des valeurs de précision moyenne (AP) est calculée sur les valeurs de rappel de 0 à 1 [80].

La formule mAP est basée sur les sous-métriques suivantes :

- Matrice de confusion
- Intersection sur union
- Rappel
- Précision

Le mAP est calculé en trouvant la précision moyenne (AP) pour chaque classe, puis en faisant la moyenne sur un certain nombre de classes.

4.8.2 Affichage de la courbe

Cette étape cruciale fournit des informations précieuses sur les performances et le comportement du modèle formé et permet de détecter toute lacune potentielle dans le processus de formation, permettant ainsi d'apporter les ajustements et améliorations nécessaires.

Implémentation & Résultats

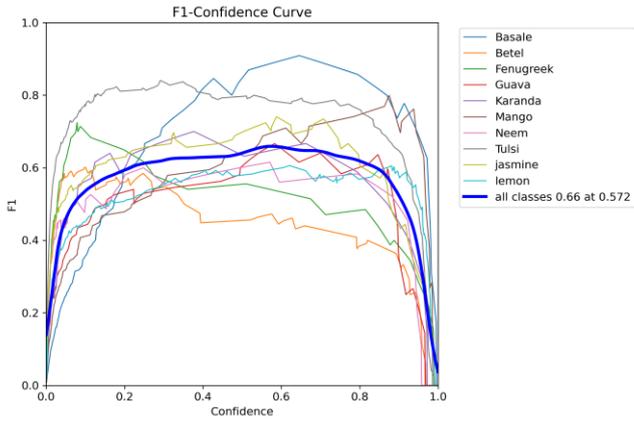


Figure 4.6 : Box F1-Curve

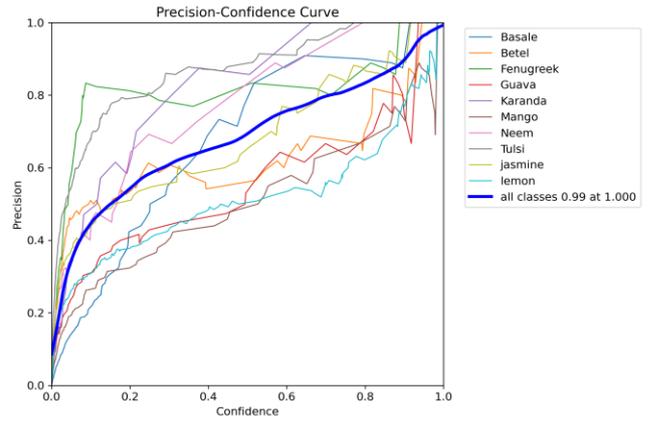


Figure 4.7 : Box P-Curve

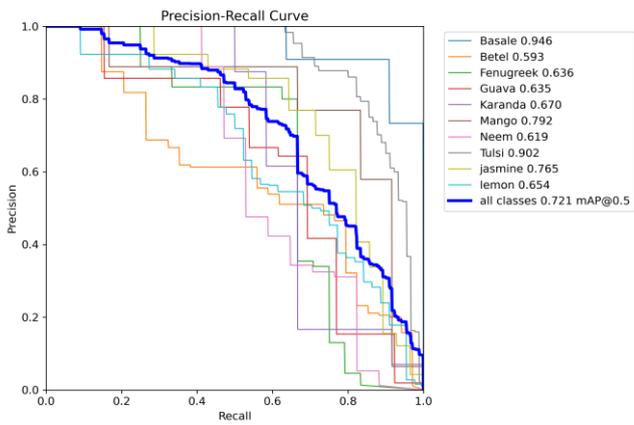


Figure 4.8 : Box PR-Curve

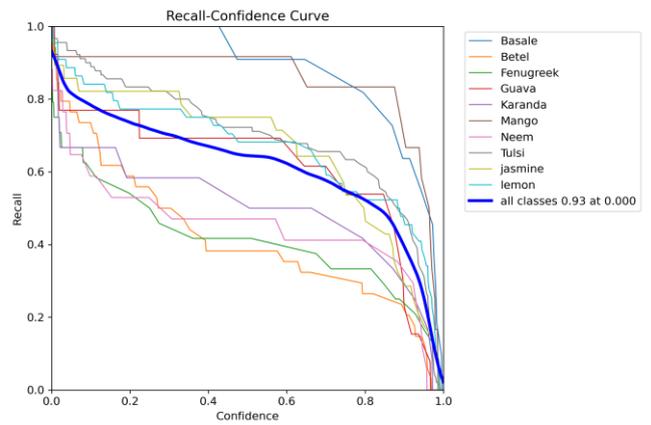


Figure 4.9 : Box R-Curve

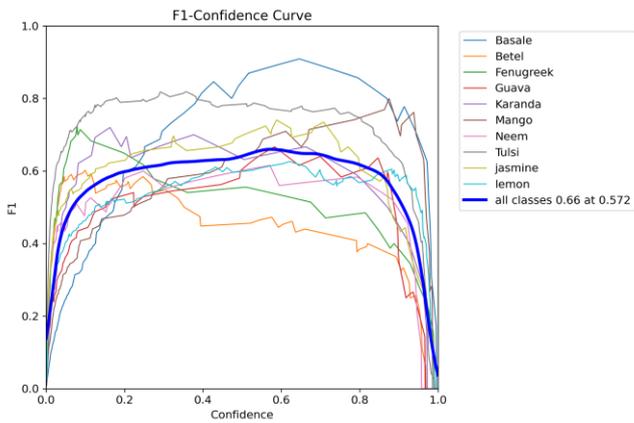


Figure 4.10 : Mask F1-Curve

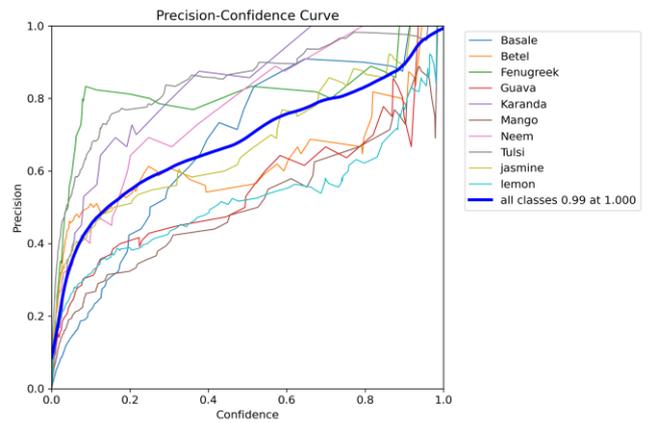


Figure 4.11 : Mask P-Curve

Implémentation & Résultats

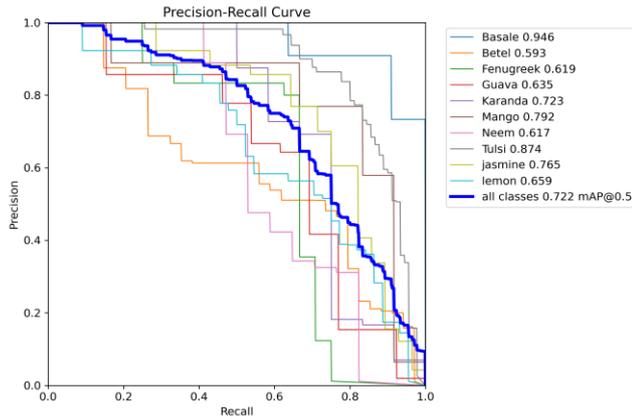


Figure 4.12 : Mask PR-Curve

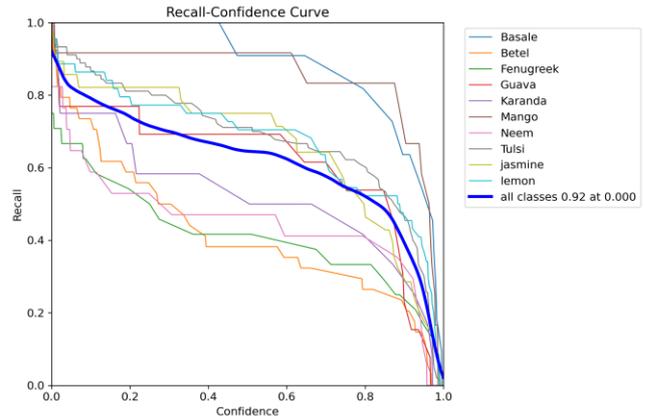


Figure 4.13 : Mask R-Curve

Courbe de confiance en matière de précision, qui est une représentation graphique des valeurs de précision à différents seuils. Cette courbe vous permet de voir comment la précision change avec le seuil. Courbe de confiance de rappel En conséquence, cette figure illustre comment la valeur de rappel change en fonction de différents seuils.

La courbe Précision-Rappel est un outil important pour évaluer l'efficacité du modèle YOLOv8seg. Cette courbe illustre le compromis entre précision et rappel, donnant un aperçu des performances du modèle. Le modèle a très bien fonctionné, avec un rappel d'environ 0,72 au niveau de confiance maximum dans toutes les classes. Un rappel aussi élevé indique l'efficacité du modèle pour identifier les plantes.

La courbe de confiance F1 est essentielle pour évaluer l'efficacité d'un modèle YOLOv8seg. Cette courbe illustre la confiance du modèle dans ses prévisions et le score F1 qui en découle, qui équilibre précision et rappel. Les performances du modèle sont impressionnantes, avec un score F1 d'environ 0,66 au niveau de confiance maximale pour toutes les classes de plantes. Ce score F1 élevé indique l'efficacité du modèle à reconnaître avec précision ces éléments essentiels.

4.8.3 Affichage de la matrice de confusion

La matrice représentée dans la figure 4.12 représente les étiquettes de test précises le long de l'axe des x et les étiquettes anticipées le long de l'axe des y. La matrice de confusion est utilisée pour évaluer avec précision les prédictions du modèle par rapport aux données réelles.

Implémentation & Résultats

La matrice est un instrument précieux pour détecter les catégories qui nuisent à l'efficacité ou à la précision du modèle. Sur la base des données fournies, il est possible de modifier délibérément les informations sur les classes problématiques et de recycler le modèle si cela est jugé nécessaire. Le processus itératif est un facteur essentiel pour améliorer la précision du modèle dans une large mesure.

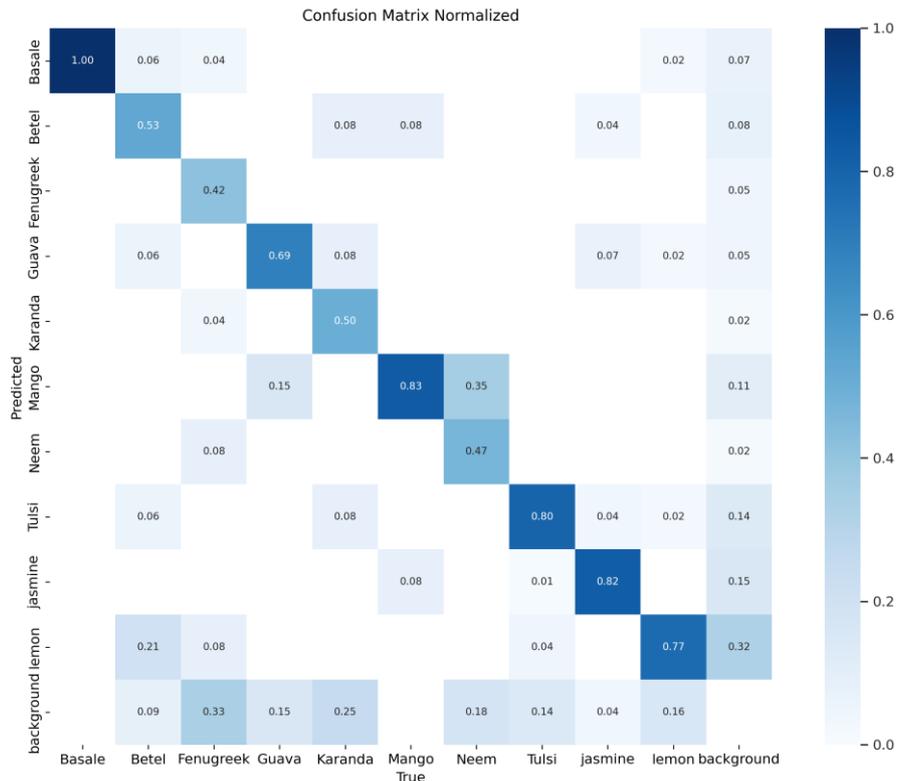


Figure 4.14 : La matrice de confusion.

Les résultats de notre modèle en ce qui concerne l'identification correcte et incorrecte des plantes médicinales sont mis en évidence dans cette matrice.

On remarque que la diagonale de la matrice a des valeurs élevées pour la majorité des marques. Cela démontre la puissance de notre modèle à détecter et à classer avec précision les plantes médicinales. Les valeurs hors diagonale, qui correspondent aux erreurs d'identification et de classification (faux positifs et faux négatifs), sont habituellement faibles, ce qui confirme la fiabilité de notre modèle pour la reconnaissance des plantes. L'efficacité de notre méthode de reconnaissance de plantes est prouvée par ces résultats encourageants, mettant en évidence le potentiel de notre modèle dans des applications concrètes.

Implémentation & Résultats

4.8.4 Résultat

L'objectif principal du processus de formation YOLO est de minimiser la fonction de perte pour améliorer les performances et la précision du modèle. Cette fonction de perte est composée de quatre types de pertes : box-loss, cls-loss, seg-loss et dfl-loss, chacune servant un objectif spécifique dans le processus de formation. À la fin de chaque période de formation, YOLO permet une analyse approfondie de ses performances grâce au calcul de deux valeurs de perte différentes. La première est connue sous le nom de perte de formation, elle sert à déterminer la fonction de coût spécifiquement pour les données de formation, et la seconde est la perte de validation, représentant la fonction de coût calculée à l'aide des données de validation.

La figure 4.15 illustre la progression de la fonction de perte sur 30 époques d'entraînement pour le modèle. Il ressort clairement du graphique que la valeur de la perte connaît une forte diminution significative de la perte en boîte, de la perte en segmentation et de la perte en dfl qui suivie d'une faible fluctuation mais finie par toujours diminuer. En ce qui concerne le cls_loss, il dimunie progressivement suivies d'un pic et atteignant une valeur de 0.5.

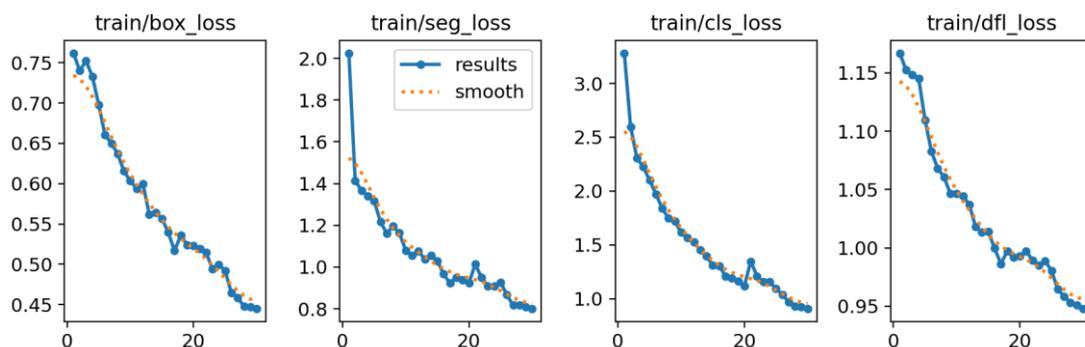


Figure 4.15 : La fonction de perte change au fil de la formation.

Les résultats de la validation sont présentés dans la figure 4.16, démontrant une petite diminution suivie d'une augmentation de la valeur de la perte box . Par la suite, il y a une diminution ultérieure suivie d'une fluctuation relativement instable caractérisée par des diminutions et des augmentations intermittentes. La perte segmentée, en revanche, présente une diminution puis une augmentation substantielle de sa valeur par la suite, il y a une forte fluctuation. En revanche, la perte dfl présente une forte fluctuation suivie d'une diminution significative jusqu'à ce qu'elle atteigne un point où il y a une fluctuation relativement instable

Implémentation & Résultats

caractérisée par des diminutions. Quant au perte cls, il présente une forte diminution ou il y a une fluctuation relativement instable caractérisée par des diminutions.

Malgré ces fluctuations, le modèle présente un niveau d'efficacité louable, démontrant sa capacité à produire des résultats favorables.

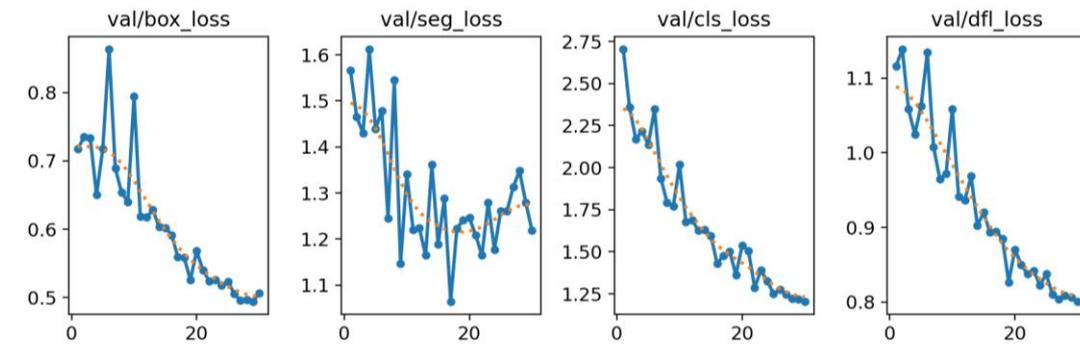


Figure 4.16 : La fonction de perte change au fil de la validation.

Cependant, s'appuyer uniquement sur la fonction de perte comme outil de surveillance pendant la formation et la validation nécessite l'ajout de mesures supplémentaires afin d'évaluer en profondeur l'efficacité et l'exactitude de la procédure de formation. Diverses méthodologies ont été développées pour évaluer les modèles YOLO, en particulier l'itération la plus récente, YOLOv8. En particulier, le score F1, la précision et le rappel sont des mesures essentielles pour évaluer l'efficacité des modèles de segmentation d'instance. Bien que ces mesures fournissent des informations pertinentes, elles ne donnent qu'une image partielle et limitée de la performance globale du modèle ; elles devraient donc être complétées par une méthode d'évaluation plus approfondie.

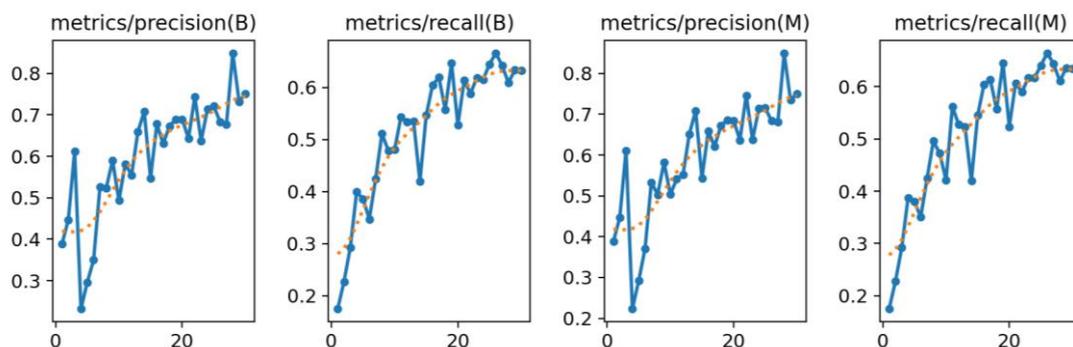


Figure 4.17 : Résultat de la précision et du rappel.

Implémentation & Résultats

La moyenne de la précision moyenne(mAP50-95) se distingue comme une meilleure statistique pour évaluer la précision d'un modèle YOLO en raison de son approche globale. Contrairement aux scores Précision, Rappel et F1, qui sont essentiels mais se concentrent sur des aspects spécifiques, le mAP50-95 fournit une évaluation complète en considérant des objets de tailles et de formes variables. Par conséquent, le mAP50-95 fournit l'indicateur le plus précis et le plus précis des performances globales du modèle.

Sur la base de notre analyse de la figure 4.18, nous avons observé une augmentation exponentielle remarquable de la précision, du rappel et du mAP50-95 au cours de la formation initiale. Cependant, montrent une amélioration significative des mesures de notation une fois la phase de formation principale commencée.

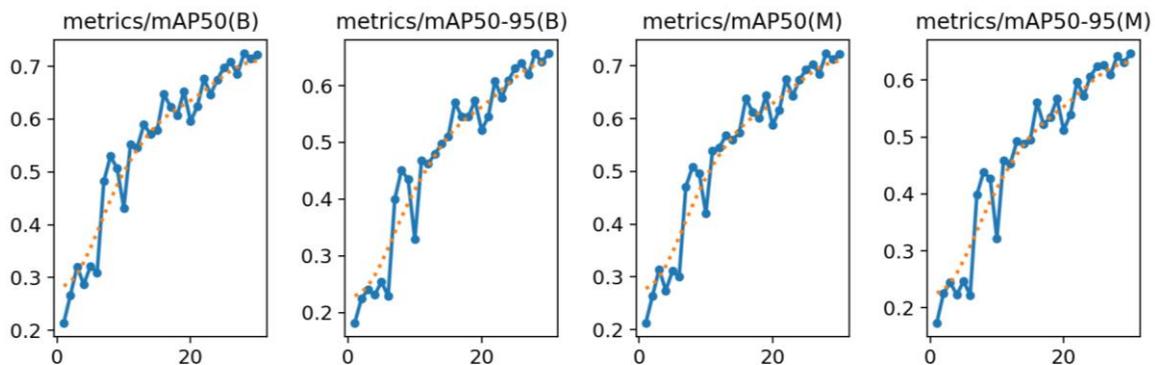


Figure 4.18 : Résultat mAP50-95.

4.8.5 Evaluation du modèle

Cette étape particulière revêt une importance significative pour déterminer l'efficacité de la formation du modèle. L'ensemble de tests d'images a réussi. Le modèle est conçu pour faire des prédictions basées sur des données et les comparer ensuite avec les étiquettes de test. Cette procédure aide en outre à identifier les cas de surajustement.

Les performances du modèle sur les images de l'ensemble de données de test peuvent être considérées comme tout à fait satisfaisantes. Cette conclusion est étayée par les exemples de prédictions présentés dans la figure 4.19, qui révèlent les meilleurs résultats obtenus par le modèle.

Implémentation & Résultats

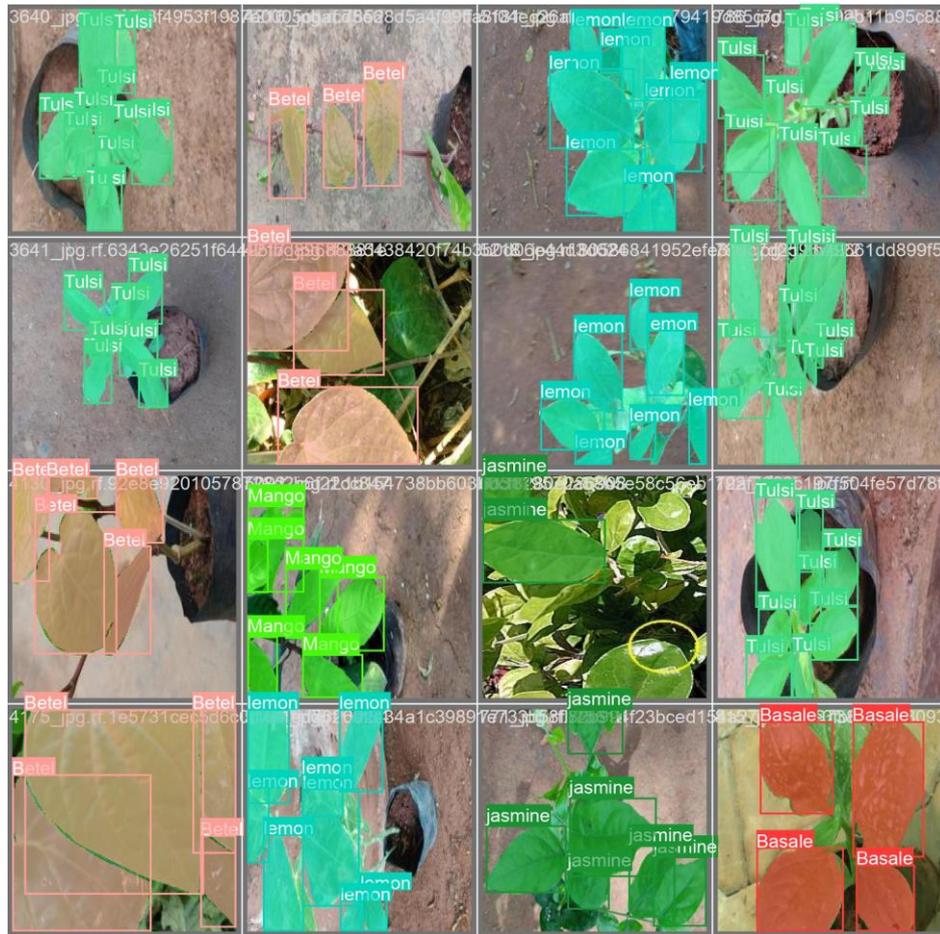


Figure 4.19 : Prédications du modèle sur l'ensemble de données de test.

Notamment, le modèle démontre un taux de rappel élevé, identifiant avec précision presque toutes les instances de segmentation avec un haut niveau de précision.

Sur la base de l'analyse complète des mesures de performance obtenues, on peut conclure avec certitude que le modèle présente une capacité remarquable à généraliser efficacement à de nouvelles données.

4.9 Comparaison avec YOLOv5 et YOLOv7

Ces modèles ont été entraînés sur la même base de données.

| Modèle | Précision | Rappel | Map0.5 |
|--------|-----------|--------|--------|
| YOLOv8 | 95.3% | 84.9% | 72% |
| YOLOv5 | 77% | 52% | 64% |
| YOLOv7 | 68% | 66% | 67% |

Table 4.1 : Comparaison avec d'autres modèles.

Implémentation & Résultats

4.10 Conclusion

Au cours de ce chapitre, nous avons exposé les logiciels et le matériel utilisés ainsi que notre modèle YOLOv8, ainsi que les différentes étapes requises et les résultats obtenus. En ce qui concerne notre modèle, il a une précision de 95,3%.

Conclusion générale

Ce travail consiste en la conception et l'implémentation d'une application mobile d'identification des plantes médicinales. Nous avons réalisé la détection de ces objets en utilisant l'architecture YOLOv8 principalement la segmentation des instances. La technologie existante combine les capacités de l'application, de la vision par ordinateur et de l'apprentissage automatique afin de détecter et d'identifier avec précision les plantes médicinales en temps réel. Le système que nous proposons vise à fournir une solution efficace et fonctionnelle en temps réel.

Dans le premier chapitre de notre mémoire, nous avons abordé les notions générales telles que l'Intelligence Artificielle, Apprentissage Automatique, Apprentissage profond et les différents travaux réalisés dans ces 2 domaines.

Dans le deuxième chapitre nous avons abordé la vision par ordinateur. Notre attention s'est particulièrement portée sur les diverses méthodes de détection et de segmentation d'images. Dans le troisième chapitre, nous avons présenté l'architecture globale du système proposé, ainsi que les différentes méthodes utilisées pour sa conception et son évaluation.

Dans le quatrième chapitre, nous avons abordé les divers outils nécessaires à la réalisation d'un tel système. De plus, nous avons présenté et discuté les résultats obtenus, accompagnés de graphiques pertinents. Pour évaluer les performances de détection des plantes médicinales, nous avons utilisé des métriques telles que la précision, le rappel et la moyenne de la Précision moyenne (mAP).

Les résultats obtenus pour le modèle d'identification des plantes sont remarquables, avec une précision de 0,93.9 et un rappel de 0,84.

Tout au long de ce travail, nous avons été confrontés à diverses difficultés et obstacles qui ont enrichi notre expérience. Parmi ceux-ci, nous avons rencontré des défis liés à la collecte de données pour notre base de données, ainsi que des contraintes liées à l'utilisation de nos ordinateurs, qui sont plus anciens par rapport aux dernières technologies. De plus, la limitation des ressources GPU dans Google Colab a également été un défi à surmonter.

Malgré ces difficultés, ce projet s'est avéré extrêmement bénéfique pour nous. Il nous a permis d'acquérir de nouvelles connaissances et d'améliorer celles que nous avons déjà acquises. En particulier, nous avons pu approfondir nos compétences dans le domaine du traitement d'images, en explorant des techniques avancées de segmentation, de détection et de reconnaissance.

Travaux futurs

A l'issue des travaux menés dans le cadre de ce mémoire, nous citons quelques perspectives qui peuvent être réalisées :

- Bien que le système actuel ait donné des résultats positifs, il existe encore de nombreuses possibilités de progrès et d'expansion futurs. En améliorant continuellement les données de formation (base de donnée des plantes algériennes) et en menant des expériences complètes à l'aide de diverses techniques, des améliorations remarquables peuvent être obtenues en termes de précision et d'efficacité du système de détection.
- Ajouter à notre système d'identification des plantes médicinales une classe de rejet car notre modèle ne traite que 10 plantes et si nous testons une plante qui ne fait pas partie de notre base de données notre modèle essayera de prédire une classe parmi les 10 et ne précise pas que la plante ne fait pas partie de notre base de données.

Ces pistes de travaux futurs illustrent le potentiel d'amélioration et de développement continu de l'application mobile et de notre système de détection des plantes Médicinales. En explorant ces possibilités, notre objectif est d'affiner les capacités de l'application mobile, de remédier aux limitations existantes et de contribuer à son progrès continu.

Glossaire

Taille du lot(batch size) La taille du lot est un hyper-paramètre qui définit le nombre d'échantillons à traiter avant de mettre à jour les paramètres internes du modèle. Cependant, augmenter la taille du modèle peut diminuer la quantité de mémoire GPU disponible pour l'entrée.

Box-loss La box-loss mesure la différence entre la boîte englobante prédite et la boîte englobante de vérité terrain.

Cls-loss La cls-loss mesure la différence entre les probabilités de classe prédites et les probabilités de classe de vérité terrain.

Dfl-loss La perte dfl mesure la différence entre la couche convolutive déformable prédite et la couche convolutive déformable de vérité terrain.

Score F1 Le score F1 est la moyenne harmonique de précision et de rappel.

Intersection sur Union (IoU) L'intersection sur Union indique le chevauchement des coordonnées de la boîte englobante prédite avec la boîte de vérité terrain. Un IoU plus élevé indique que les coordonnées prévues de la boîte englobante ressemblent étroitement aux coordonnées de la boîte de vérité terrain.

La moyenne de la précision moyenne (mAP50-95) Il s'agit d'une métrique qui fournit une image plus complète des performances du modèle en mesurant la précision moyenne moyenne sur une plage de seuils d'intersection sur l'union (IoU) de 0,5 à 0,95. Cette métrique prend en compte le rappel et la précision du modèle à différents seuils IoU.

Précision La précision mesure le pourcentage de détections correctes parmi toutes les détections.

Rappel Le rappel mesure le pourcentage de détections correctes parmi tous les objets de vérité terrain.

Seg-loss Fait référence à la perte de segmentation. Il s'agit d'une mesure ou d'une fonction qui quantifie la dissemblance entre la carte ou le masque de segmentation prédit et la carte de segmentation de vérité terrain.

Bibliographie

Bibliographies

- [1] Abdul Mueed Hafiz and Ghulam Mohiuddin Bhat. A survey on instance segmentation: state of the art. *International Journal of Multimedia Information Retrieval*, 9(3):171–189, September 2020.
- [2] Agus Ambarwari, et al. Plant species identification based on leaf venation features using SVM. *TELKOMNIKA Telecommunication, Computing, Electronics and Control* Vol. 18, No. 2, April 2020.
- [3] ahim Azadnia, et al. AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling. *Agronomy* 2022,2,novembre 2022.
- [4] Akshit Mehra. Evolution of yolo object detection model from v5 to v8, Mar 2023.
- [5] Alexander Kirillov, et al. Panoptic Segmentation. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9396–9405, Long Beach, CA, USA, June 2019.IEEE.
- [6] Alexandru Munteanu and Marian Neagul. Semantic Segmentation of Vegetation in Remote Sensing Imagery Using Deep Learning. September 2022.
- [7] Alexy Bhowmick, et al. Non-parametric scene parsing: Label transfer methods and datasets. *Computer Vision and Image Understanding*, 219:103418, 2022.
- [8] Andra Petrovai and Sergiu Nedevschi. Fast Panoptic Segmentation with SoftAttention Embeddings. *Sensors (Basel, Switzerland)*, 22(3):783, January 2022.
- [9] Ani Aggarwal. YOLO Explained, January 2021.
- [10] Ashish Jaiswal and Aditya Dwivedi. Python: The Versatile Language – Recent Trends in Programming Languages. 08:2021, May 2021.
- [11] Belkacem KOUNINEF (2) - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Les-domaines-de-IIA-Source_fig1_371533311,2018
- [12] Bus service quality prediction and attribute ranking using probabilistic neural network and adaptive neuro fuzzy inference system - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Architecture-of-probabilistic-neural-network-PNN_fig1_304038706 [ccessed 4 Jun, 2024]
- [13] Daniel Bolya, et al. YOLACT++: Better Real-time Instance Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2):1108–1121, February 2022. arXiv:1912.06218 [cs, eess].

Bibliographie

- [14] Deep neural network for fringe pattern filtering and normalisation - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-U-net-and-V-net-architectures-look-similar-at-block-level-However-the-number-of_fig1_333815410 [ccessed 4 Jun, 2024]
- [15] Efficient Reliability Analysis of Structures Using Symbiotic Organisms Search-Based Active Learning Support Vector Machine - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Illustration-of-linear-SVM-Classifier-separating-the-two-classes-Illustration-of-linear_fig1_359803757 [ccessed 9 Jun, 2024]
- [16] Evan Tarver. Market segmentation: Definition, example, types, benefits, Mar 2023.
- [17] Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Typical-process-of-segmentation-with-Deep-Learning-A-Convolutional-Neural-Network-CNN_fig5_334498427 [ccessed 11 May, 2024]
- [18] Hind Alaskar and Tanzila Saba. Machine Learning and Deep Learning: A Comparative Review. pages 143–150. April 2021.
- [19] Hongshan Yu, et al. Methods and datasets on semantic segmentation: A review. *Neurocomputing*, 304:82–103, 2018.
- [20] Investigations of Object Detection in Images/Videos Using Various Deep Learning Techniques and Embedded Platforms—A Comprehensive Review - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/RCNN-architecture-17_fig4_341099304 [ccessed 4 Jun, 2024]
- [21] James G Shanahan and Liang Dai. Introduction to computer vision and real time deep learning-based object detection. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3523–3524, 2020.
- [22] Kannan Kattari, et al. Identification of Medical Plants Using Logistic Decision Regression Algorithm. Conference paper, janvier 2023.
- [23] Kashvi Taunk, et al. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. Pages 1255–1260, May 2019.
- [24] Laura von Rueden, et al. Informed machine learning—a taxonomy and survey of integrating knowledge into learning systems. arXiv preprint arXiv:1903.12394, 2019.

Bibliographie

- [25] Md. Khairul Islam, Sk Md Masudul Ahsan, Sultana Habiba. Bangladeshi Plant Leaf Classification and Recognition Using YOLO Neural Network. Conference paper, decembre 2019.
- [26] Mostafa Mehdipour Ghazi, et al. Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, 26, April 2017.
- [27] Natalie Kudan. How to label images for semantic segmentation, Jan 2023.
- [28] Nilesh Barla. Panoptic segmentation: Definition, datasets & tutorial [2023], April 24 2023.
- [29] Omar Elharrouss, et al. Panoptic Segmentation: A Review. November 2021.
- [30] Osmar Luiz Ferreira de Carvalho, et al. Panoptic Segmentation Meets Remote Sensing. *Remote Sensing*, 14(4):965, January 2022. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- [31] Owais A. Malik, Nazrul Ismail, Burhan R. Hussein Umar Yahya. Automated Real-Time Identification of Medicinal Plants Species in Natural Environment Using Deep Learning Models—A Case Study from Borneo Region. 27, juillet 2022.
- [32] Prasad Ramesh. Customer segmentation: Explained with 6 examples and 8 models, June 8 2022.
- [33] Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051 [ccesed 4 Jun, 2024]
- [34] Puneet Jindal. YOLOv8 Is Here and It Gets Better!, January 2023.
- [35] Robert G. de Luna, et al. Identification of Philippine Herbal Medicine Plant Leaf Using Artificial Neural Network. Conference paper, decembre 2017.
- [36] Roza Dastres and Mohsen Soori. Artificial Neural Network Systems. *International Journal of Imaging and Robotics*, 21:13–25, March 2021.
- [37] Shahadat Uddin, et al. Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19, December 2019.
- [38] Shan Suthaharan. Support Vector Machine, pages 207–235. 01 2016.
- [39] Shervin Minaee, et al. Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7):3523–3542, July

Bibliographie

2022. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [40] Shubham Mahajan, et al. Plant Recognition Using Morphological Feature Extraction and Transfer Learning over SVM and AdaBoost.22,fevrier 2021.
- [41] Sik-Ho Tsang. Review: Fcn - fully convolutional network (semantic segmentation), Mar 2019.
- [42] Stephen Gang Wu, et al. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network.2007
- [43] Tiago Carneiro, et al. Performance analysis of google colab as a tool for accelerating deep learning applications. IEEE Access, 6:61677–61685, 2018.
- [44] Tingting Yang , et al. An Approach for Plant Leaf Image Segmentation Based on YOLOV8 and the Improved DEEPLABV3+.Plants 2023,29,septembre 2023.
- [45] Tsung-Yi Lin, et al. Feature Pyramid Networks for Object Detection. 19 Avril 2017.
- [46] Visualizing and Understanding Convolutional Networks for Semantic Segmentation Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Architecture-of-the-DeconvNet-The-encoder-network-is-based-on-VGG16-followed-by-a_fig1_367180367 [ccesed 11 May, 2024]
- [47] Wan Muhammad Arief Hilman Bin Wan Zefri, Sharifalillah Nordin. Plant Recognition System Using Convolutional Neural Network.IOP Conference Series: Earth and Environmental Science,23,novembre 2021
- [48] Waseem Khan. Image Segmentation Techniques: A Survey. Journal of Image and Graphics, pages 166–170, 2014.
- [49] Wei Mao, et al. Can we cover navigational perception needs of the visually impaired by panoptic segmentation? July 2020.
- [50] Wirdayanti, et al. Identification of Herbal Plants Using Morphology Method and K-Nearest Neighbour Algorithm (KNN). IOP Conference Series: Earth and Environmental Science,2022
- [51] Xiang Yue, et al. Improved YOLOv8-Seg Network for Instance Segmentation of Healthy and Diseased Tomato Plants in the Growth Stage.Agriculture 2023,21,aout 2023.

Bibliographie

- [52] Yousef Sharrab, et al. Plants Recognition Using Deep Learning. The International Conference on Multimedia Computing, Networking and Applications (MCNA2023)At: VALENCIA, SPAIN,juin 2023.
- [53] Yu Sun, et al. Deep learning for plant identification in natural environment. Computational intelligence and neuroscience, 2017, 2017
- [54] Yuliia Kniazieva. What is semantic segmentation in computer vision?, September 22 2022.
- [55] AHEAD of AI. Comprendre les LLM d'encodeur et de décodeurUnderstanding Encoder And Decoder LLMs. <https://magazine.sebastianraschka.com/i/128842439/encoder-and-decoder-style-transformers> , Juin 2023.
- [56] Ashish Patel. What is Object Detection?, June 2020.
- [57] Boryl. Firebase. <https://www.boryl.fr/glossaire/firebase/> ,Decembre 2020.
- [58] Docker. [https://hub.docker.com/r/ultralytics/ultralytics?](https://hub.docker.com/r/ultralytics/ultralytics?fbclid=IwAR2Idgp_U7nHZpyDWu0fXxjSNrS1Inyh5A8eWpF5rZlSYubQ4yNerbX21r8) fbclid=IwAR2Idgp_U7nHZpyDWu0fXxjSNrS1Inyh5A8eWpF5rZlSYubQ4yNerbX21r8.Retrieved April 27, 2023.
- [59] FasterCapital.CollecteDeDonnéesEtPrétraitement.<https://fastercapital.com/fr/sujet/collectededonn%C3%A9setpr%C3%A9traitement.html#:~:text=La%20collecte%20de%20donn%C3%A9es%20est,des%20donn%C3%A9es%20pour%20l'analyse> ,20, Mars 2024
- [60] GeeksforGeeks.Dart tutorial. <https://www.geeksforgeeks.org/dart-tutorial/> ,Avril 2024
- [61] GeeksforGeeks.KNN algorithm. <https://www.geeksforgeeks.org/k-nearest-neighbours/> ,Janvier 2024.
- [62] Hmrishav Bandyopadhyay. The definitive guide to instance segmentation [+v7 tutorial]. <https://www.v7labs.com/blog/instance-segmentation-guide>, April 24 2023.
- [63] IBM Cloud Education. What are convolutional neural networks? <https://www.ibm.com/cloud/learn/convolutional-neural-networks>, October 2020. Accessed: 2022-6-14.
- [64] Ikomia. What is YOLO? An In-Depth Introduction to Object Detection in Computer Vision. <https://www.ikomia.ai/blog/what-is-yolo-introduction-object-detection-computer-vision> , 23,octobre 2023.

Bibliographie

- [65] Ishan Gaba. What is github and how to use it? Simplilearn.com; Simplilearn, June 2020.
- [66] LEDATASCIENTIST. Qu'est ce que le deep learning (Apprentissage profond) ?. <https://ledatascientist.com/deep-learning-apprentissage-profond/> , 2024
- [67] LePont. Computer vision : tout connaître sur cette technique de l'IA. <https://www.lepont-learning.com/fr/computer-vision/> ,2024
- [68] LinkedIn.Object Detetction from traditional techniques to modern deep learning approaches. <https://www.linkedin.com/pulse/object-detection-from-traditional-techniques-modern-samer/> ,Aout 2023.
- [69] Medium. Probablist neural network. <https://makquel.medium.com/probabilistic-neural-networks-in-a-nutshell-28d56470e9c3> ,Juin 2021
- [70] Medium. What is convolutionnal neural network. <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5> ,Fevrier 2023
- [71] MINDY SUPPORT. What is image segmentation. <https://mindy-support.com/news-post/what-is-image-segmentation-the-basics-and-key-techniques/> ,Octobre 2022.
- [72] Papers. Reseau de neurones convolutifs. <https://paperswithcode.com/method/spp-ne> ,Juin 2014.
- [73] Papers.Detectiond'objet.<https://paperswithcode.com/method/retinanet#:~:text=RetinaNet%20is%20a%20one%2Dstage,learning%20on%20hard%20negative%20examples> , Decembre 2019.
- [74] SlidePlayer.Intelligenc artificielle apprentissage supervisee. <https://slideplayer.fr/slide/15169377/> , 2019
- [55] UX24/7. What usability heuristics can do for you? <https://ux247.com/usability-heuristics-can/> , April 2019.
- [76] V7 Editorial Team. YOLO: Algorithm for Object Detection Explained [+Examples], January 2023.
- [77] V7. The Ultimate Guide to Object Detection. <https://www.v7labs.com/blog/object-detection-guide> , 2021

Bibliographie