



*République Algérienne Démocratique et
Populaire*

*Ministère de l'enseignement supérieur et de la
recherche scientifique*

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

Département : Mathématiques et Informatique

*Mémoire de fin d'étude
Pour l'obtention du diplôme de MASTER*

Domaine : Mathématiques et Informatique

Filière : Informatique

Option : Système et multimédia

***JA pour la sécurité alimentaire
Cas d'étude : Les maladies des plantes***

Thème Présenté Par : Saker Oumaima

Devant le jury

<i>Dr. SOUAFI MOHAMMED SALAH</i>	<i>MCB</i>	<i>Université Larbi Tébessi</i>	<i>Président</i>
<i>Dr. GHREIB NAWEL</i>	<i>MCB</i>	<i>Université Larbi Tébessi</i>	<i>Examinateur</i>
<i>Pr. LAOUER MOUHAMED RIDDA</i>	<i>Pr</i>	<i>Université Larbi Tébessi</i>	<i>Encadreur</i>
<i>M.LLE. TOUATI Hamad ZAINEB</i>		<i>Université Larbi Tébessi</i>	<i>CO.Encadreur</i>

Date de soutenance : 07/06/2023



Remerciment

Tout d'abord, je tiens à exprimer ma gratitude envers le Tout-Puissant pour m'avoir accordé la force et la patience nécessaires pour mener à bien ce travail de thèse

Ce projet est le résultat d'un parcours long et difficile durant lequel j'ai bénéficié du soutien, de l'encouragement et de l'encadrement de plusieurs personnes. Nous souhaitons exprimer notre gratitude profonde et sincère envers toutes ces personnes.

*Je suis profondément reconnaissante envers mon encadreur, le **Pr. LAOUAR MOHAMED RIDDA**, d'avoir accepté de me superviser tout au long de mon travail. Ses précieux conseils et orientations ont été d'une grande aide*

*Ma gratitude va à mon encadrante, **Mlle Zaineb Touati Hamad**, pour son soutien tout au long de mon travail. Sa patience, son expérience, ses conseils précieux et ses efforts ont été d'une aide inestimable.*

Enfin je tiens également à exprimer ma gratitude envers les membres du jury,pour avoir accepté de discuter de notre projet de fin d'étude et nous faire profiter de leur expérience et de leurs précieux conseils

Dédicace

Je suis honoré de dédier cet humble travail à ceux qui m'ont
encouragé et soutenu.

Mon cher père qui a toujours été à mes côtés à tout moment.

Ma chère mère, qui me donne toujours espoir dans la vie.

Mes sœurs pour leur soutien.

Toute ma famille

Mon meilleur et plus cher ami qui m'a encouragé tout au long de
mon parcours : *CHAIMA*

Que Dieu me les garde.

Sommaire

Table des matières

Remerciement.....	II
Dédicace	III
Sommaire.....	IV
Liste des figures.....	VII
Liste des tableaux	VIII
Abstract.....	X
Introduction générales	A

CHAPITRE 1 : SECURITE ALIMENTAIRE ET LES MALADIES DES PLANTES

1 Introduction.....	2
2 Sécurité alimentaire mondiale.....	2
3 Les défis mondiaux affectés par la sécurité alimentaire	3
4 La sécurité alimentaire	3
4.1 L'objectif de la sécurité alimentaire.....	4
4.2 Dimension principales de la sécurité alimentaire	4
4.2.1.La disponibilité des aliments	4
4.2.2L'accès économique et physique des aliments.....	4
4.2.3L'utilisation des aliments	5
4.3 Les étapes clés pour atteindre la sécurité alimentaire	5
5 Impact des maladies sur les aliments	5
6 Définition des maladies et ravageurs des plantes	6
7 Les maladies des plantes	6
7.1 Les Types de maladies des plantes	7
8 Taxonomie des maladies des cultures et de leurs symptômes	8
9 L'effet de facteur météorologique sur les maladies Agricole.....	10
10 Méthodes de détection des maladies des plantes.....	10
11 Les maladies de tomates	11
12 Conclusion.....	16

CHAPITRE 2 : Intelligence Artificielle

1 Introduction.....	18
2 Intelligence artificielle	18
3 Apprentissage automatique	19
4 Apprentissage en profondeur	20
5 Apprentissage en profondeur VS apprentissage automatique	21

6 Types d'apprentissages	21
6.1 Apprentissage supervisé.....	21
6.2 Apprentissage non supervisé.....	22
6.3 Apprentissage par renforcement	22
7 Apprentissage Métrique	23
7.1 L'apprentissage ponctuel (One shot Learning)	23
7.2 Les modèles d'apprentissage métrique	24
7.2.1.Siamese Networks	24
7.2.2.Contrastive Loss	25
7.2.3.Triplet Networks.....	26
8 IA pour les maladies des plants	27
8.1 Les jeux de données	27
8.2 Les méthodes traditionnelles pour le traitement des maladies des plantes	28
8.3 Apprentissage en profondeur pour les maladies des plantes.....	28
8.3.1 Apprentissage de la représentation	31
8.3.1.1 Classification	31
8.3.1.2 Détection.....	33
8.3.1.3 Segmentation	34
8.3.2Apprentissage Métrique.....	35
9 Les défis	37
10 Synthèse	40
11 Conclusion.....	41

CHAPITRE 3: CONTRIBUTION ET EXPERIMENTATION

1 Introduction.....	43
2 Les systèmes intelligents pour traiter les maladies des plantes	43
3 Jeux de données	44
4 Architecture de système	47
4.1 Prétraitement	48
4.2 Modélisation	48
4.2.1 Modèle :CNN (from scratch).....	49
5 Implémentation	51
5.1 Le langage de programmation	51
5.1.1. Les librairies et les bibliothèques Python	51
5.1.2. Plateforme de développement	52

6	Expérimentation et Résultat.....	53
6.1	Les métriques évalué.....	53
6.2	Résultat	54
6.3	TEST.....	55
6.4	Discussion.....	56
7	Conclusion	56

CONCLUSION GENERALE ET PERCPECTIVE

Liste des figures

Figure 1: Les dimensions principales de la sécurité alimentaire.....	4
Figure 2: maladies des plantes	6
Figure 3 : type des maladies agricole	7
Figure 4: Différents types d'agents pathogènes : virale, fongique et bactériennes [19].....	8
Figure 5: Le triangle des maladies des plantes.....	10
Figure 6: les méthodes de détection des maladies des plantes.....	11
Figure 7: Différence entre IA et machine apprentissage automatique et profond	18
Figure 8: Les types d'apprentissage supervisé.....	22
Figure 9: Les différents modèles utilisé dans Metric Learning.....	24
Figure 10: le modèle de réseau siamese	25
Figure 11: la perte contrastive entre les paires d'entrées	26
Figure 12: Un modèle de triplet	27
Figure 13 : Les méthodes de traitement des maladies et ravageurs de plantes basée sur le deep learning.....	30
Figure 14 : Le traitement des maladies de plantes	31
Figure 15: Les méthodes de traitement les maladies des plantes basé sur l'apprentissage métrique.....	43
Figure 16: Conception globale	47
Figure 17: Division de dataset.....	48
Figure 18: La figure illustre l'architecture SNN.....	49
Figure 19: Le Modèle CNN	49
Figure 20: Les couches de modèle	50
Figure 21: Editeur de code et bibliothèques Python utilisés	51
Figure 22: graphe de précision et de perte	54
Figure 23: Etats de sortie de système	55

Liste des tableaux

Tableau 1: Classification de certaines maladies foliaires avec leurs symptômes	9
Tableau 2 : Description des maladies de la tomate incluses dans l'étude, définissant chaque maladie et décrivant ses symptômes	13
Tableau 3: Comparaison entre l'apprentissage supervisé et non-supervisé	20
Tableau 4: Les caractéristiques principales qui distinguent l'apprentissage en profondeur de l'apprentissage automatique.....	21
Tableau 5: Les ensembles de données auto-collectés sur les maladies de plantes (tomate)	28
Tableau 6 : Approches basées sur l'apprentissage profond pour détecter et classer les feuilles des plantes de tomates	36
Tableau 7: Description de la base de données.....	45
Tableau 8: Etude comparative.....	56

Résumé

La sécurité alimentaire est une préoccupation majeure à l'échelle mondiale, et l'intelligence artificielle (IA) joue un rôle de plus en plus crucial dans ce domaine. Dans ce contexte, la reconnaissance automatique des maladies des plantes, telle que la détection des maladies de la tomate à partir d'images de feuilles, revêt une importance particulière pour prévenir les pertes de récolte et garantir une production alimentaire durable. Les méthodes de reconnaissance des maladies de la tomate basées sur l'apprentissage en profond ont montré des résultats prometteurs, mais elles sont confrontées à des défis majeurs. Elles nécessitent souvent de vastes ensembles de données étiquetées, ce qui peut être coûteux et fastidieux à obtenir.

De plus, ces modèles d'apprentissage en profondeur traditionnels consomment beaucoup de mémoire et de stockage en raison du grand nombre de paramètres qu'ils utilisent. Dans ce travail, un cadre léger basé sur un réseau siamois a été développé pour la reconnaissance automatique des maladies des feuilles de tomate. Ce cadre léger a atteint une précision remarquable sur un sous-ensemble de données de tomates provenant de l'ensemble de données PlantVillage. Les résultats expérimentaux démontrent l'efficacité de ce cadre dans la gestion de données déséquilibrées et de petite taille. De plus, le réseau profond intégré à ce cadre se distingue par sa légèreté ce qui le rend bien inférieur aux réseaux profonds légers existants.

Mot clés : intelligence artificielle, apprentissage métrique, réseau neuronal siamois, plantvillage, maladies des tomates

Abstract

Food Security is a major global concern, and artificial intelligence (AI) is playing an increasingly crucial role in this area. In this context, automatic recognition of plant diseases, such as tomato disease detection from leaf images, is of particular importance to prevent crop losses and ensure sustainable food production. Deep learning-based tomato disease recognition methods have shown promising results, but they face major challenges. They often require large sets of labeled data, which can be expensive and time-consuming to obtain.

Moreover, these traditional deep learning models consume a lot of memory and storage due to the large number of parameters they use. In this work a lightweight framework based on a Siamese network was developed for the automatic recognition of tomato leaf diseases. This lightweight framework achieved a remarkable accuracy on a subset of tomato data from the PlantVillage dataset. The experimental results demonstrate the effectiveness of this framework in handling unbalanced and small data. In addition, the backbone deep network integrated into this framework is distinguished by its lightness, which makes it far inferior to existing light deep networks.

Keywords : artificial intelligence, metric learning, Siamese neural network, plantvillage, tomato diseases

ملخص

تعتبر سلامة الغذاء مصدر قلق عالمي رئيسي، ويلعب الذكاء الاصطناعي (AI) دورًا مهمًا بشكل متزايد في هذا المجال. في هذا السياق، يعتبر التعرف التلقائي على أمراض النبات، مثل الكشف عن أمراض الطماطم من صور الأوراق، ذا أهمية خاصة لمنع خسائر المحاصيل وضمان الإنتاج الغذائي المستدام. أظهرت طرق التعرف على أمراض الطماطم القائمة على التعلم العميق نتائج واعدة، لكنها تواجه تحديات كبيرة. غالبًا ما تتطلب مجموعات كبيرة من البيانات المصنفة، والتي قد تكون باهظة الثمن وتستغرق وقتًا طويلًا للحصول عليها.

علاوة على ذلك، تستهلك نماذج التعلم العميق التقليدية هذه قدرًا كبيرًا من الذاكرة والتخزين نظرًا للعدد الكبير من المعلمات التي تستخدمها. في هذا العمل، تم تطوير إطار خفيف الوزن يعتمد على شبكة سيامية للتعرف التلقائي على أمراض أوراق الطماطم. حقق إطار العمل خفيف الوزن هذا دقة ملحوظة تبلغ على مجموعة فرعية من بيانات الطماطم من مجموعة بيانات PlantVillage. تظهر النتائج التجريبية فعالية هذا الإطار في التعامل مع البيانات غير المتوازنة والصغيرة. بالإضافة إلى ذلك، تتميز الشبكة الأساسية العميقة المدمجة في هذا الإطار بخفتها، مما يجعلها أدنى بكثير من الشبكات العميقة الخفيفة الموجودة.

الكلمات المفتاحية: ذكاء اصطناعي، تعلم متري، شبكة عصبية سيامية، PlantVillage، أمراض طماطم

Liste des abréviations

- **IA** : intelligence artificielle
- **SNN** : réseau neuronal siamois
- **CNN** : réseau neuronal convontionel
- **SSD** : Détecteur MultiBox à un coup (Single Shot MultiBox Detector)
- **YOLO** : vous ne regardez qu'une seule fois (you only look once)
- **FCN** : réseaux entièrement convolutifs (fully convolutional networks)
- **R-CNN** : Réseau neuronal convolutif régional (Region-based Convolutional Neural Network)
- **VGG** : Groupe de géométrie visuelle (Visual Geometry Group)

Introduction générale

Introduction générales

Les pays en développement sont susceptibles d'être les plus durement touchés par les conséquences du changement climatique, tels que l'augmentation des températures, les changements dans les régimes de précipitations, l'élévation du niveau de la mer et l'augmentation des phénomènes météorologiques extrêmes. Tous ces effets représentent des risques pour l'agriculture et la disponibilité de nourriture. L'agriculture est particulièrement confrontée à un défi majeur : les épidémies non contrôlées entraînent souvent d'énormes pertes économiques en raison de l'incapacité de traiter les maladies au bon moment.

C'est pourquoi il est essentiel de mettre en place un système de prédiction des maladies dès l'apparition des premiers symptômes sur les feuilles des plantes. Cette approche permettrait d'anticiper les épidémies et d'intervenir de manière préventive, en prenant des mesures appropriées pour contrôler les maladies avant qu'elles ne se propagent et ne causent des dommages importants aux cultures.

Existe-t-il une solution intelligente pour atténuer les effets des maladies agricoles et limiter les pertes économiques causées par l'incapacité de traiter les épidémies de manière opportune?

Dans le secteur de l'agriculture, l'utilisation de l'apprentissage automatique, en particulier de la Machine Learning (ML), suscite un intérêt croissant pour alimenter les algorithmes intelligents. Dans le cadre de ce mémoire, notre objectif est de concevoir et mettre en œuvre un système d'apprentissage automatique qui permet de détecter et de classer automatiquement les feuilles des plantes de tomate. Dans ce contexte, l'utilisation de la Metric Learning permet d'apprendre une métrique de similarité entre les images, ce qui permet d'identifier les maladies même lorsque l'apparence de la plante affectée présente des différences légères. Cette approche offre l'avantage de fournir une évaluation plus nuancée du niveau de maladie présent, permettant ainsi une détection plus précise et une meilleure compréhension des variations dans les symptômes des maladies.

Ce mémoire est structuré en trois chapitres. Le premier chapitre aborde le sujet de la sécurité alimentaire et des maladies des plantes. Il met en évidence l'importance de ces problématiques et leurs impacts sur l'approvisionnement alimentaire.

Le deuxième chapitre se concentre sur l'intelligence artificielle et présente un état de l'art sur les maladies affectant spécifiquement la tomate. Ce chapitre explore les avancées dans ce domaine et offre une synthèse des travaux de recherche associés.

Le troisième chapitre présente la conception de notre système, en mettant en évidence l'architecture SNN (réseau neuronal siamois) que nous avons proposée. Nous décrivons également les outils d'implémentation utilisés et la plateforme sur laquelle notre travail a été réalisé. Enfin, nous exposons les résultats que nous avons obtenus dans le cadre de notre recherche.

Chapitre 01

**Sécurité alimentaire et les
maladies des plants**

1 Introduction

Les plantes représentent environ 80 % de l'alimentation humaine. Même ainsi, ils sont nécessaires à la sécurité alimentaire ou à l'accès continu à une alimentation adéquate, accessible, sûre et saine pour que nous puissions tous mener une vie active et saine. La salubrité des aliments est menacée par les ravageurs et les maladies des végétaux, car ils peuvent endommager les cultures, réduisant ainsi la disponibilité et l'accès à la nourriture et augmentant le coût des aliments.

La sapidité des aliments peut également être affectée négativement par les ravageurs et les maladies des plantes, entraînant des changements dans les préférences alimentaires traditionnelles des populations. Aujourd'hui, les ravageurs et les maladies des végétaux constituent une menace plus grande que jamais pour la sécurité alimentaire.

2 Sécurité alimentaire mondiale

En 2012, Ban Ki-moon, le Secrétaire général des Nations Unies (ONU), a initié le Défi faim zéro. Ce défi vise à promouvoir une alimentation durable tout au long de la chaîne alimentaire, de la production à la consommation, dans le but d'éradiquer la pauvreté, d'augmenter la productivité agricole et les revenus, de prévenir le gaspillage alimentaire, de garantir l'accès à une alimentation suffisante pour tous, ainsi que de promouvoir une alimentation saine pour tous. L'objectif est de mettre fin à la malnutrition d'ici 2030

La sécurité alimentaire est menacée par les pénuries alimentaires à mesure que la population augmente. La population devrait augmenter de 8 milliard d'ici 2050, passant de 7,3 milliards actuellement (UNFPA) à 9 milliards en 2050, qui souffriront de l'insécurité alimentaire.[1]

Cette analyse montre que le gaspillage alimentaire a été atteint à un niveau inacceptable. La plupart des pays en développement qui manquent d'installations financières, managériales et techniques, de refroidissement et de stockage sont confrontés au problème. [2]

La FAO rapporte que sur l'ensemble de la population, 60% d'entre eux sont des femmes qui ont faim et 5 millions d'enfants meurent de malnutrition à l'âge de cinq ans[3]. Ces personnes sont celles qui ne reçoivent pas suffisamment de nourriture pour répondre aux besoins alimentaires, est passée de 23,3% en 1990-1992 à 12,9% en 2015. La fourchette moyenne de 795 millions de personne.[4]

L'agriculture familiale est une excellente solution pour réduire la faim dans le monde. Dans le monde, pas moins de 500 millions de personnes possèdent des exploitations agricoles familiales. Ces agriculteurs comprennent parfaitement l'histoire de la terre, ce dont nous avons besoin et ce que la terre peut produire. La culture de base du Brésil a augmenté de 40 % et celle des exploitations familiales américaines de 84 %. Cela comprend les fermes maraîchères, les fermes alimentaires, les pêcheurs et les personnes qui récoltent les produits.[5]

La pauvreté est un autre problème qui conduit à la faim chronique. Pour faire face à ce type de problèmes, l'ONU a lancé un défi Faim zéro. En 1961, l'organisme nommé Programme alimentaire mondial (PAM) aide à fournir des plans de nutrition et de sécurité alimentaire. Il aide également à réduire les problèmes causés par la sous-nutrition et contribue également à réduire les problèmes de santé chroniques qui nécessitent un traitement urgent pour sauver la vie des personnes.[6]

À l'avenir, avec la croissance simultanée de la population, la production nationale devrait augmenter. En outre, des techniques modernes de sécurité alimentaire auraient dû être utilisées pour protéger les ressources naturelles pour l'agriculture et la durabilité agricole. Par conséquent, assurez l'efficacité des ressources à chaque étape en formant et en encourageant les parties prenantes. Et le partage d'informations peut contribuer à une bonne production et à moins de gaspillage alimentaire.

3 Les défis mondiaux affectés par la sécurité alimentaire

Les défis qui affectent la sécurité alimentaire [7]

- Les changements climatiques et les facteurs économiques.
- Distribution de nourriture
- La production de nourriture
- La santé humaine
- Les moyens de subsistance
- Le pouvoir d'achat de nourriture

Autres facteurs qui affectent l'insécurité alimentaire mondiale

- Augmentation de la population
- Urbanisation
- Exigences alimentaires
- Ressources naturelles
- Infrastructures et communication.

4 La sécurité alimentaire

La notion de sécurité alimentaire a été introduite à l'occasion de la Conférence mondiale sur l'alimentation organisée par l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) en 1974[8].

La sécurité alimentaire repose sur deux aspects essentiels : la disponibilité alimentaire et le droit à l'alimentation. La disponibilité alimentaire concerne l'approvisionnement en nourriture au niveau national et international. Le droit à l'alimentation fait référence à la capacité des individus et des ménages à se procurer de la nourriture, ce qui inclut également ceux qui n'ont pas les ressources financières nécessaires pour acheter des aliments nutritifs [9].

La sécurité alimentaire a été définie par la FAO en 1996 et affinée en 2001 comme "*Lorsque chaque individu dispose d'un accès physique, social et économique à une quantité adéquate d'aliments sûrs, nutritifs et adaptés à leurs besoins et préférences, cela leur permet de mener une vie active et saine*". " [10]

La sécurité alimentaire peut être définie de différentes manières selon les auteurs, mais elles partagent toutes un sens commun. En résumé, nous pouvons définir la sécurité

alimentaire comme étant l'état où chaque individu a accès à une quantité suffisante d'aliments nutritifs, lui permettant ainsi de mener une vie saine et active.

4.1 L'objectif de la sécurité alimentaire

L'objectif ultime de la sécurité alimentaire est de veiller à ce que toutes les personnes aient accès à la nourriture dont elles ont besoin pour mener une vie saine et productive, tout en promouvant des systèmes alimentaires durables et résilients capables de s'adapter à l'évolution des circonstances et de répondre aux besoins des générations futures.

4.2 Dimension principales de la sécurité alimentaire

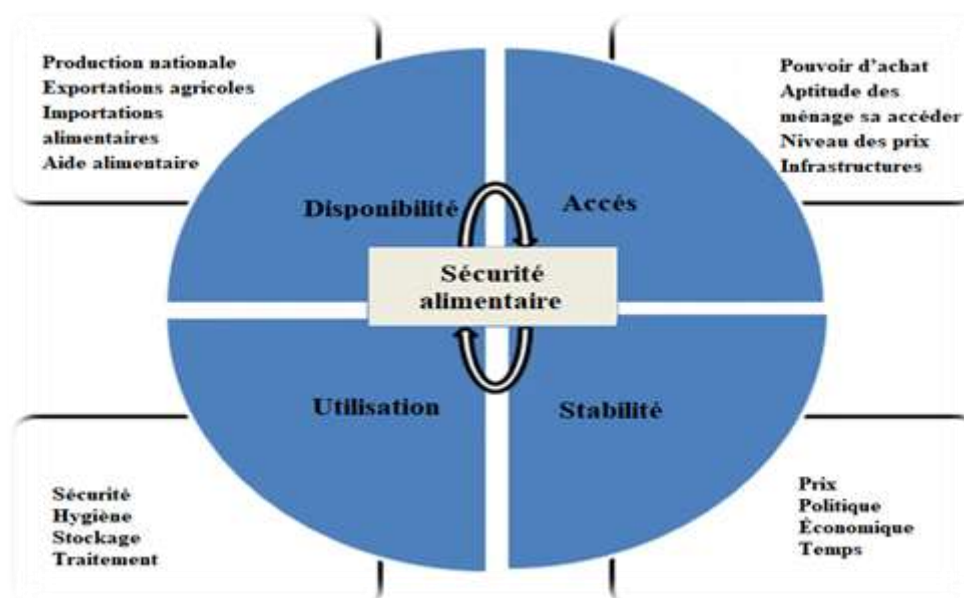


Figure 1: Les dimensions principales de la sécurité alimentaire

4.2.1. La disponibilité des aliments

C'est l'existence d'une nourriture physique qui comprend suffisamment de nourriture devrait être disponible pour que les gens satisfassent pleinement leurs besoins alimentaires, c'est-à-dire, par une bonne production et une bonne capacité d'importation de denrées alimentaires, se traduit par une disponibilité accrue.[11]

4.2.2 L'accès économique et physique des aliments

Il devrait y avoir suffisamment de ressources disponibles pour acheter la qualité de la nourriture avec une quantité suffisante qui inclura économiquement et physiquement les gens peuvent accéder à la nourriture comme si les gens avaient un bon revenu et avaient facilement accès à la commodité du transport, puis la capacité d'acheter des aliments nutritifs Sera augmentée. [11]

4.2.3 L'utilisation des aliments

Entre en scène lorsque la nourriture est disponible et accessible aux personnes, il est alors temps d'utiliser les aliments avec sécurité, hygiène, stockage et transformation des aliments, c'est-à-dire, protégeant les aliments des bactéries, bien préparés avec hygiène et bon usage du réfrigérateur permettront une meilleure utilisation des aliments.[11]

4.2.4 La stabilité des trois dimensions dans le temps

Les ménages bénéficient d'un approvisionnement alimentaire constant tout au long de l'année. Le changement du prix des aliments, les facteurs politiques, les problèmes économiques et les changements climatiques affectent la stabilité alimentaire, c'est-à-dire qu'en diminuant le gaspillage alimentaire et en contrôlant la population, la stabilité alimentaire sera atteinte.[11]

4.3 Les étapes clés pour atteindre la sécurité alimentaire

Pour assurer la sécurité alimentaire, plusieurs étapes clés doivent être prises en compte. [12]

Tout d'abord, il est primordial d'augmenter la productivité agricole en utilisant des techniques améliorées, des semences de qualité et des engrais appropriés, tout en investissant dans la recherche et le développement de nouvelles technologies. En parallèle, il est essentiel d'améliorer l'accès à la nourriture en renforçant les infrastructures de transport et de stockage des aliments, et en soutenant financièrement les ménages à faible revenu.

Réduire le gaspillage alimentaire revêt également une grande importance, impliquant l'amélioration des pratiques de stockage et de manipulation des denrées. La promotion de systèmes alimentaires durables est un autre aspect crucial, englobant la préservation des ressources naturelles, la réduction des émissions de gaz à effet de serre et l'assurance d'une rémunération équitable pour les producteurs.

Enfin, il est impératif de s'attaquer aux causes profondes de la faim en luttant contre la pauvreté, en favorisant l'accès à l'éducation et en promouvant l'égalité des sexes. Il est également essentiel de protéger les systèmes alimentaires contre les perturbations telles que le changement climatique et les conflits.

5 Impact des maladies sur les aliments

Les maladies agricoles peuvent affecter directement la productivité et la disponibilité des aliments. Cependant, en détectant les cultures / maladies des plantes à un stade précoce, cela aide à produire une bonne production alimentaire.[13]

La maladie est essentiellement définie comme chaque fois qu'une plante a changé son état de normal à malade, ce qui peut directement interrompre les fonctions vitales de la plante. La plupart des espèces végétales peuvent souffrir de maladies. Ces maladies sont causées par certains agents pathogènes appelés champignons, virus et bactéries. Certains des agents

vivants sont également responsables de maladies des plantes telles que les insectes, les rongeurs, les puces, les mouches et les cafards, etc.

6 Définition des maladies et ravageurs des plantes

Les maladies et les ravageurs des plantes sont un type de catastrophes naturelles qui affectent la croissance normale des plantes et même provoquent la mort des plantes pendant tout le processus de croissance des plantes, du développement des graines aux semis et à la croissance des semis.[14]



Figure 2: maladies des plantes [15]

7 Les maladies des plantes

Les maladies des plantes sont causées par des agents similaires à ceux qui affectent les humains et les animaux. En général, deux types de facteurs ont un impact sur les plantes : les facteurs animés, c'est-à-dire vivants, et les facteurs abiotiques, qui sont non vivants[16].

- Les agents biologiques : Ces agents comprennent les insectes, les bactéries, les champignons et les virus, qui sont responsables de diverses maladies chez les plantes.
- Les facteurs abiotiques : Ce sont des influences environnementales telles que les fluctuations rapides de température, l'excès d'humidité, les carences en éléments nutritifs, le pH du sol et les conditions d'humidité élevée, qui peuvent également affecter la santé des plantes.

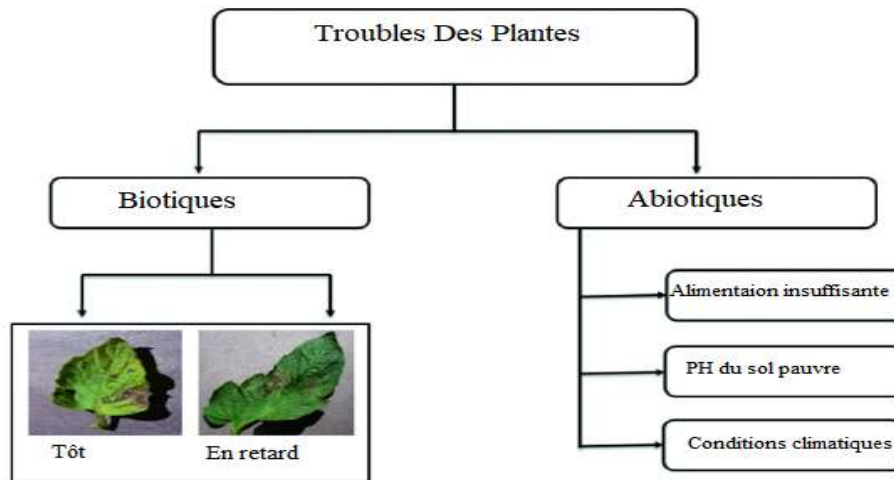


Figure 3 : type des maladies agricole [17]

7.1 Les Types de maladies des plantes

- **Maladie virale**

Parmi toutes les maladies des plantes, celles causées par l'infection sont difficiles à identifier et à diagnostiquer ; de plus, ces symptômes sont confondus avec des signes de carence nutritionnelle ou de blessure, car il n'existe aucun indicateur préconçu pouvant être surveillé en permanence. Les aleurodes, les cicadelles, les pucerons et les insectes rampants du concombre sont des vecteurs réguliers de maladies virales.[18]

- **Maladies bactériennes**

Les agents pathogènes provoquent des maladies graves chez les légumes. Ils ne pénètrent pas directement dans la végétation, mais plutôt par des blessures ou des ouvertures dans la culture. Les dommages aux cultures résultent de divers agents pathogènes, insectes et outils agricoles lors de tâches telles que la cueillette et l'élagage.[18]

- **Maladie fongique**

Les maladies foliaires sont généralement provoquées par des champignons tels que le mildiou, l'anthracnose et l'oïdium. Elles se manifestent d'abord sur les feuilles inférieures âgées, où l'on observe des taches gris-vert ou des zones imbibées d'eau. Avec le temps, ces taches s'assombrissent et favorisent le développement des champignons responsables de la maladie.[18]

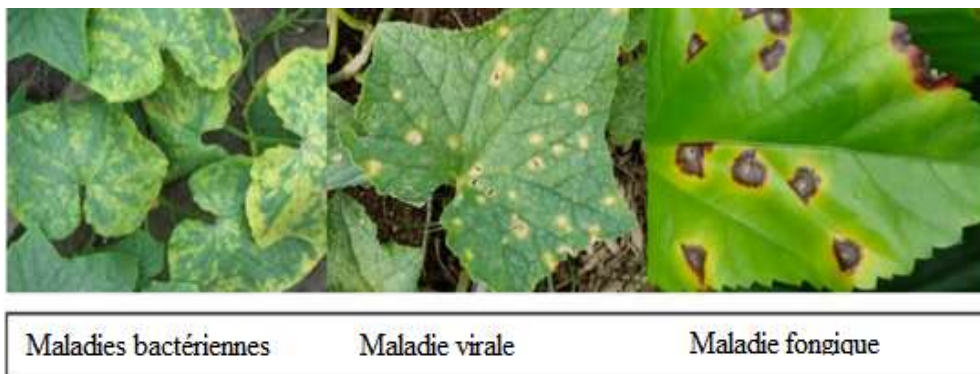


Figure 4: Différents types d'agents pathogènes : virale, fongique et bactériennes [19]

8 Taxonomie des maladies des cultures et de leurs symptômes

Les cultures sont souvent vulnérables aux maladies, un phénomène naturel qui peut affecter la qualité et la quantité des récoltes si des mesures de prévention ne sont pas prises en temps opportun.

Les pathogènes responsables des maladies peuvent être des organismes autotrophes qui se nourrissent de tissus vivants ou des saprophytes qui se nourrissent de tissus morts. Les symptômes de la maladie sont visibles sur les feuilles et peuvent affecter leur couleur, leur forme et leur texture. Ainsi, la surveillance visuelle des feuilles est utile pour détecter les maladies telles que le mildiou, la rouille et l'oïdium.[20]

Tableau 1: Classification de certaines maladies foliaires avec leurs symptômes [19]

Feuille de plante	Maladies	Symptômes	Catégorie d'agent pathogène
Riz	• Tache brune/Bipolaris oryzae	• Centre blanchâtre gris	• Champignons
	• Blast leaf/Pyricularia oryzae Cavara	• Un brun foncé irrégulier	
Coton	• Mildiou aréolé/Cercospora	• jaune clair avec des marges brun foncé	• Virus
	• Tache foliaire/tache alternarienne	• Tache bronzée brune	• Champignons
	• Brûlure bactérienne/Xanthomonas campestris	• Auréole jaunâtre vert	• Bactérien
Tomate	• Alternariose/Alternaria tomatophila	• Tache en anneau sombre autour de lui jaune	• Champignons
	• Mildiou/Phytophthora infestans	• La tache sombre se développe rapidement	• Champignons
	• Oïdium/Leveillula taurica	• Feuille frisée et jaunâtre	• Champignons
	• Virus de la chlorose infectieuse des boucles jaunes/de la tomate	• Trempé dans l'eau cernée d'un halo jaune	• Virus

9 L'effet de facteur météorologique sur les maladies Agricole

Des études suggèrent que le changement climatique peut perturber les taux et les stades de développement des agents pathogènes, ainsi que la résistance des plantes hôtes, ce qui modifie les interactions physiologiques entre l'hôte et le pathogène. De plus, une utilisation inappropriée de pesticides peut entraîner une résistance à long terme des agents pathogènes, compromettant ainsi les efforts de lutte contre les maladies.[16]

En présence de conditions climatiques favorables, les agents pathogènes se multiplient rapidement, créant une interaction complexe entre l'hôte, le climat et les agents pathogènes. Cette dynamique est souvent représentée par le concept du "triangle des maladies", comme illustré dans la figure ci-dessous.

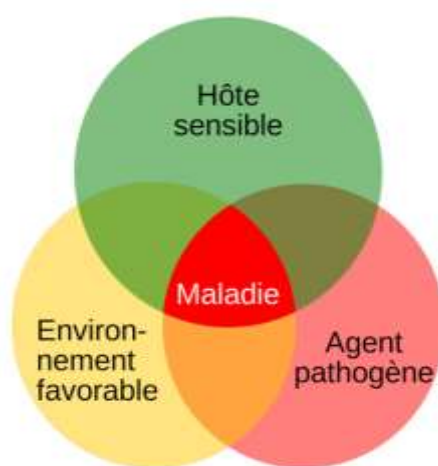


Figure 5: Le triangle des maladies des plantes [21]

La quatrième dimension, à savoir le temps, est devenu un élément clé dans l'étude des épidémies végétales. Selon la pyramide des maladies, les épidémies se produisent lorsque la perturbation entre les différents éléments tels que les plantes, le climat, les agents pathogènes et la météo persiste pendant des périodes prolongées, allant de quelques heures à plusieurs jours. La prédiction des systèmes épidémiques des plantes repose sur la compréhension des interactions entre ces éléments de la pyramide. En particulier, la surveillance des conditions climatiques peut aider à déterminer si les pathogènes sont dans une phase active ou passive de développement, permettant ainsi de prévenir l'apparition de maladies.

10 Méthodes de détection des maladies des plantes

Après avoir pris des mesures pour traiter les symptômes d'une maladie végétale, il est essentiel de vérifier la présence de cette maladie dans les plantes en utilisant des techniques de détection.

À l'heure actuelle, deux principales techniques sont couramment utilisées à cet effet : l'ELISA, qui se base sur l'utilisation de protéines produites par l'agent pathogène, et la PCR, qui repose sur des séquences spécifiques d'ADN du pathogène. Cependant, il existe une

demande croissante pour une méthode de détection sélective permettant une identification rapide des maladies des plantes. Les techniques de détection peuvent être classées en techniques directes ou indirectes, et la figure ci-dessous récapitule ces différentes méthodes.[22]

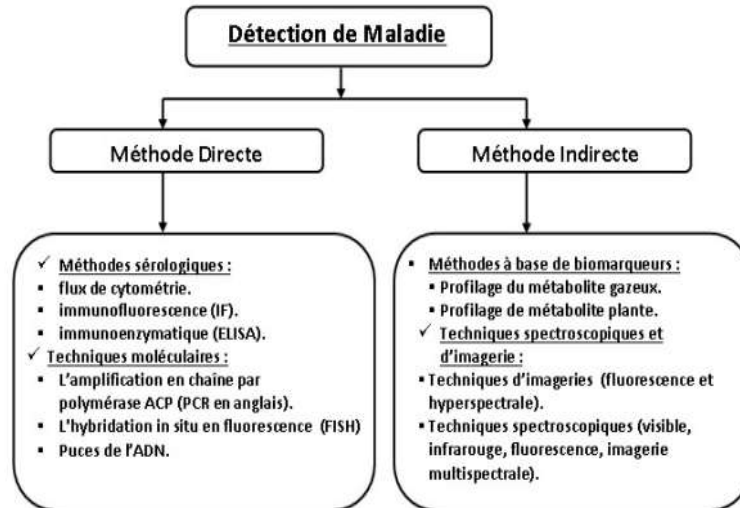


Figure 6: les méthodes de détection des maladies des plantes_[22]

Il existe deux approches distinctes pour détecter les maladies affectant les plantes. La première méthode consiste à utiliser des techniques d'imagerie pour identifier les symptômes de la maladie sur les plantes. La seconde approche repose sur l'utilisation de biomarqueurs appelés métabolites organiques volatils pour détecter les maladies.

Ces deux approches sont particulièrement intéressantes car elles peuvent également être intégrées à des véhicules, ce qui permet une surveillance rapide, fiable et en temps réel des cultures. La détection précoce et précise des maladies végétales revêt une grande importance économique, pour la production agricole et pour la santé des plantes, offrant ainsi de nombreux avantages. Des études sur le terrain sont menées de manière cohérente afin de collecter des données spectrales dans des conditions réelles.

- Les techniques moléculaires telles qu'ELISA et PCR sont couramment utilisées pour la détection des maladies végétales. D'autres techniques moléculaires comprennent IF, cytométrie de flux, FISH et les puces d'ADN. [22]
- Les techniques spectroscopiques et d'imagerie sont des méthodes de surveillance uniques qui peuvent détecter les maladies dès les premiers symptômes sur les plantes. Ces techniques peuvent être intégrées dans des véhicules agricoles autonomes pour fournir des informations de détection précoce des maladies et contrôler leur propagation.[22]

11 Les maladies de tomates

La tomate (*Solanum lycopersicum* L.) est une légume largement cultivée et consommée dans le monde entier. Cependant, les maladies qui provoquent les cultures de tomates peuvent

limiter leur croissance et leur qualité, représentant ainsi une partie importante des coûts de production.

La tomate est l'une des cultures les plus largement cultivées et consommées dans le monde elle est une culture de base dans de nombreux pays et joue un rôle vital dans la sécurité alimentaire.

Les maladies de la tomate peuvent entraîner des pertes économiques importantes pour les agriculteurs et l'industrie agricole.

Les maladies affectant les plants de tomates peuvent entraîner une réduction de la production agricole, des fruits de moindre qualité et une pénurie alimentaire accrue. Le développement de méthodes efficaces de détection et de gestion des maladies peut aider à assurer un approvisionnement stable en tomates, contribuant ainsi à la sécurité alimentaire.




Les plants de tomates sont sensibles à diverses maladies, notamment les infections fongiques, bactériennes et virales. Des maladies telles que l'alternariose, le mildiou, la tache bactérienne et le virus de la mosaïque de la tomate sont répandues et peuvent causer de graves dommages aux cultures de tomates. En vous concentrant sur les maladies de la tomate, vous pouvez relever un défi agricole important.




Les pommes de terre sont également une culture vivrière importante, mais elles ne sont pas aussi sensibles aux maladies que les tomates. En effet, les pommes de terre sont des racines, tandis que les tomates sont des fruits. Les plantes-racines sont généralement moins sensibles aux maladies que les cultures fruitières car elles sont souterraines et protégées des éléments.



Les maladies de la tomate présentent souvent des symptômes visibles sur les feuilles, les tiges et les fruits. Cette caractéristique rend les maladies de la tomate adaptées à la détection basée sur l'image à l'aide d'algorithmes d'apprentissage profond. En tirant parti des techniques de vision par ordinateur, vous pouvez former des modèles pour identifier et classer différents types de maladies de la tomate en fonction de leur apparence visuelle.

Les ensembles de données sur les maladies de la tomate sont relativement abondants et accessibles au public, ce qui facilite la collecte de données de formation pour les modèles d'apprentissage en profondeur. Plusieurs projets de recherche et organisations ont compilé des ensembles de données contenant des images de plants de tomates sains et malades, permettant le développement et l'évaluation de modèles robustes de détection des maladies.

Tableau 2 : Description des maladies de la tomate incluses dans l'étude, définissant chaque maladie et décrivant ses symptômes

Classe	Description	Images
Point cible	<ul style="list-style-type: none"> • une maladie fongique qui affecte les plantes de tomates. Elle est causée par le champignon <i>Corynespora cassiicola</i> et se caractérise par l'apparition de taches circulaires sur les feuilles de la plante. • Ces taches peuvent avoir un centre brun foncé et un bord jaune clair, ce qui leur donne l'aspect d'une cible. 	
Tétranyques à deux points	<ul style="list-style-type: none"> • le tétranyque à deux points est un petit acarien qui peut causer des dommages aux plantes de tomates en se nourrissant de la sève des feuilles. • Les symptômes de l'infestation comprennent des feuilles tachetées et pâles, qui peuvent tomber prématurément de la plante 	
Tache septorienne	<ul style="list-style-type: none"> • le nom de tache foliaire septorienne, • est une maladie fongique courante qui affecte les plantes de tomate. • Les symptômes de la tache septorienne commencent généralement par l'apparition de petites taches noires ou brunes sur les feuilles de la plante. • Ces taches peuvent avoir des centres plus clairs et des bords 	

	<p>foncés, et peuvent se développer en une forme de cercle.</p>	
<p>Moisissure de feuille</p>	<ul style="list-style-type: none"> est une maladie fongique courante qui peut affecter les feuilles de nombreuses plantes, y compris les tomates. Les symptômes de la moisissure de feuille peuvent varier, mais les signes les plus courants incluent des taches blanches ou grises sur les feuilles, qui se développent en un revêtement de moisissure blanc, gris ou brun. 	
<p>Mildiou (early blight)</p>	<ul style="list-style-type: none"> cette maladie est causée par un champignon appelé <i>Phytophthora infestans</i>. Elle se manifeste par des taches sombres sur les feuilles, qui se développent ensuite en des taches brunâtres, molles et friables. Les feuilles utilisées finissent par se flétrir et mourir. Le mildiou peut également affecter les fruits de la tomate, causant leur pourriture. 	
<p>Mildiou (late blight)</p>	<ul style="list-style-type: none"> Les symptômes de la late blight incluent des lésions brunes à noires sur les feuilles, qui peuvent se développer rapidement pour couvrir la plante entière. Les symptômes de l'early blight incluent des taches brunes sur les feuilles inférieures de la plante, qui peuvent se développer en lésions circulaires plus grandes avec des centres bruns et des 	

	<p>bords concentriques.</p>	
<p>Tache bactérienne</p>	<ul style="list-style-type: none"> • Les taches foliaires bactériennes sont causées par des bactéries qui infectent les feuilles de la tomate. • Les feuilles touchées auront des taches brunes à noires avec des centres nécrotiques. 	
<p>Virus de l'enroulement des feuilles jaunes</p>	<ul style="list-style-type: none"> • Le virus de l'enroulement des feuilles jaunes (ou TYLCV, pour Tomato Yellow Leaf Curl Virus en anglais) • est une affection virale qui touche spécifiquement les plantes de tomates. Elle est provoquée par un virus appartenant à la famille des Geminiviridae et est transmise par des insectes tels que la mouche blanche. • Les symptômes du TYLCV comprennent l'enroulement des feuilles de la plante de tomate vers le haut et leur jaunissement, ainsi que la réduction de la taille des feuilles et des fruits. 	

12 Conclusion

En conclusion, la sécurité sanitaire des aliments est un problème majeur dans le monde et les maladies des plantes ont un impact significatif sur la production alimentaire et la sécurité alimentaire. Les maladies des plantes peuvent réduire la quantité et la qualité des récoltes, augmenter les coûts de production et réduire les revenus des agriculteurs, ainsi qu'entraîner des pertes économiques importantes.

Par conséquent, il est nécessaire de développer des stratégies efficaces pour prévenir et contrôler les maladies des plantes en utilisant des méthodes avancées.

Chapitre 02

Intelligence Artificielle

1 Introduction

L'intelligence artificielle (IA) est de plus en plus utilisée pour relever les défis de la sécurité alimentaire. L'IA est utilisée pour développer des techniques d'agriculture de précision qui peuvent aider les agriculteurs à optimiser les rendements des cultures et à réduire les déchets et aussi utilisée pour optimiser la gestion de la chaîne d'approvisionnement et réduire le gaspillage alimentaire.

L'IA peut être utilisée pour surveiller la croissance et la santé des cultures, et aider à identifier et à prévenir les maladies et les ravageurs des cultures il peut être utilisée pour développer des stratégies d'adaptation au changement climatique pour l'agriculture.

Dans ce chapitre ont introduit la technologie IA et les travaux connexes partagés dans notre domaine.

2 Intelligence artificielle

L'intelligence artificielle est un domaine scientifique qui a émergé dans les années 50 et qui se concentre sur l'utilisation de techniques pour permettre aux machines d'accomplir des tâches et de résoudre des problèmes qui étaient traditionnellement réservés aux humains ou à certains animaux. [23]

Les objectifs de l'IA sont de reproduire, imiter ou simuler les capacités associées à l'intelligence humaine, notamment la capacité de percevoir l'environnement, d'interagir avec lui, de communiquer grâce à un langage, de raisonner et de résoudre des problèmes, de planifier des actions pour atteindre un but, de représenter la connaissance et de l'appliquer au moment opportun.

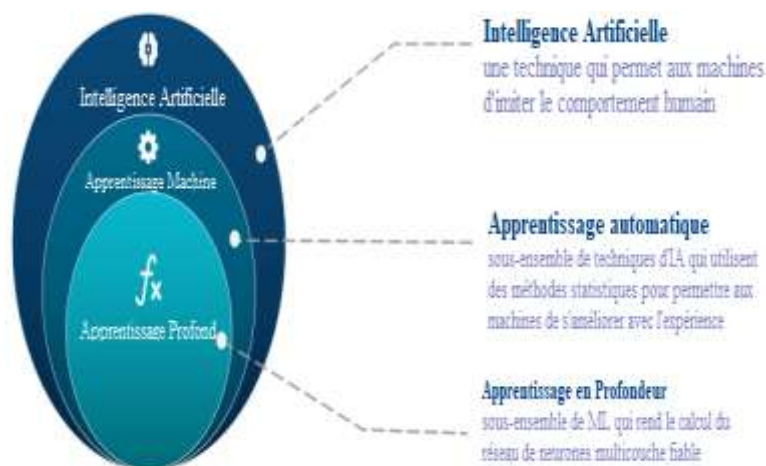


Figure 7: Différence entre IA et machine apprentissage automatique et profond[24]

Comme mentionné précédemment, le domaine de l'intelligence artificielle englobe un large éventail de disciplines qui intègrent les sciences cognitives, les mathématiques, l'électronique et l'informatique, entre autres.

Au fil du temps, le domaine de l'intelligence artificielle a connu diverses évolutions. Dans les années 80, les "systèmes experts" étaient particulièrement populaires. Cette approche

mettait l'accent sur la capacité des machines à reproduire un raisonnement logique afin de résoudre des problèmes spécifiques dans des domaines particuliers.[25]

Les programmes développés à cette époque étaient conçus pour exploiter des règles logiques préétablies par des experts humains et stockées dans des bases de connaissances. Cependant, cette approche présentait des limites lorsqu'il s'agissait de traiter des problèmes complexes impliquant de multiples variables. Une alternative consistait à s'inspirer du fonctionnement du cerveau humain pour concevoir des modèles capables de s'adapter à leur environnement. C'est dans cette optique que le domaine du Machine Learning, et plus spécifiquement du Deep Learning, a émergé.

3 Apprentissage automatique

L'apprentissage automatique, une branche de l'IA, permet aux machines d'apprendre et d'améliorer progressivement leurs performances sur une tâche spécifique en utilisant des données, sans qu'il soit nécessaire de les programmer explicitement pour résoudre cette tâche. Les tâches d'apprentissage automatique sont divisées en deux catégories principales : l'apprentissage supervisé et l'apprentissage non supervisé, qui dépendent du signal d'apprentissage que le système utilise pour apprendre. [24]

La classification est un concept fondamental de l'apprentissage automatique et fait référence au processus de catégorisation ou d'attribution d'étiquettes aux données d'entrée en fonction de leurs caractéristiques. Cela implique la formation d'un modèle sur un ensemble de données étiqueté, où chaque instance de données est associée à une classe ou une catégorie spécifique. L'objectif est de développer un modèle capable de classer avec précision les nouvelles instances de données invisibles dans les catégories appropriées.

La classification est largement utilisée dans divers domaines, notamment la reconnaissance d'images, le traitement du langage naturel, l'analyse des sentiments, la détection des fraudes et le diagnostic médical. Il permet une prise de décision automatisée et une catégorisation des données en fonction de modèles appris, permettant des applications dans les tâches de prédiction, de recommandation et de classification.

Les développeurs et chercheurs en IA ont développé une gamme d'algorithmes pour répondre à différentes problématiques. Parmi les plus connus, on peut citer :

- ✓ La méthode des k plus proches voisins, qui relève de l'apprentissage supervisé et de la classification.
- ✓ Les machines à vecteur de support (SVM), qui sont également utilisées en apprentissage supervisé pour la régression ou la classification.
- ✓ Les arbres de décision et les forêts aléatoires (Random Forest), qui sont des techniques d'apprentissage supervisé permettant de prendre en compte l'impact des décisions.
- ✓ Les réseaux de neurones, qui peuvent être utilisés tant en apprentissage supervisé qu'en apprentissage non supervisé.

Tableau 3: Comparaison entre l'apprentissage supervisé et non-supervisé [26]

	Apprentissage supervisé (Classification)	Apprentissage non-supervisé (Clustering)
Données d'entrée	Utilise les données connues et étiquetées comme entrées	Données inconnues en entrée
Complexité informatique	Très complexe	Moins de complexité informatique
Temps réel	Utilise l'analyse hors ligne	Utilise l'analyse en temps réel des données
Sous-domaines	Classification et régression	Exploitation de règles de clustering et d'association
Précision	Produit des résultats précis	Génère des résultats modérés
Nombre de classes	Nombre de classes connues	Le nombre de classes n'est pas connu

4 Apprentissage en profondeur

« L'apprentissage en profondeur est une méthode d'apprentissage automatique qui utilise des architectures hiérarchiques composées de plusieurs couches de traitement informatique itératif. Cette technique combine des algorithmes d'apprentissage supervisé et non supervisé pour des tâches d'analyse et de classification » [27]

L'apprentissage en profondeur implique le calcul des caractéristiques hiérarchiques des paramètres des réseaux de neurones artificiels pour les données d'entrée sous forme de vecteurs. Les méthodes d'apprentissage en profondeur comprennent les réseaux de neurones, les modèles probabilistes hiérarchiques, ainsi qu'une variété d'algorithmes d'apprentissage supervisés et non supervisés. Cette famille de méthodes continue de s'étendre et d'évoluer au fil du temps. [27]

5 Apprentissage en profondeur VS apprentissage automatique

Il existe deux caractéristiques principales qui distinguent l'apprentissage en profondeur de l'apprentissage automatique qui sont :

Tableau 4: Les caractéristiques principales qui distinguent l'apprentissage en profondeur de l'apprentissage automatique

	Extraction de caractéristiques	Performance
Apprentissage en profondeur	La plupart des fonctionnalités de l'application est requise par un expert puis codé manuellement par domaine et le type de données.	les résultats prévalent en fonction de la quantité de données augmente.
Apprentissage automatique	Dont les algorithmes tentent de connaître fonctionnalité de haut niveau des données	Lorsque les données sont petites, les performances de les algorithmes d'apprentissage en profondeur donnent de mauvais résultats car il nécessite une grande quantité de données pour le comprendre parfaitement.

6 Types d'apprentissages

6.1 Apprentissage supervisé

L'apprentissage supervisé est la forme la plus courante d'apprentissage automatique. Son but est de permettre au système de généraliser ce qu'il a appris pour des données inconnues en lui fournissant des exemples étiquetés.

Lorsque des données d'entraînement sont étiquetées, le système peut évaluer ses performances en comparant ses prédictions aux résultats connus, ce qui lui permet de mettre à

jour et d'améliorer le modèle. Après l'entraînement, une portion des données étiquetées peut être réservée à des fins de validation pour évaluer l'efficacité du modèle.

Les algorithmes de classification supervisée les plus connus sont les SVM, les arbres de décision, la régression logistique, les forêts aléatoires. [28]

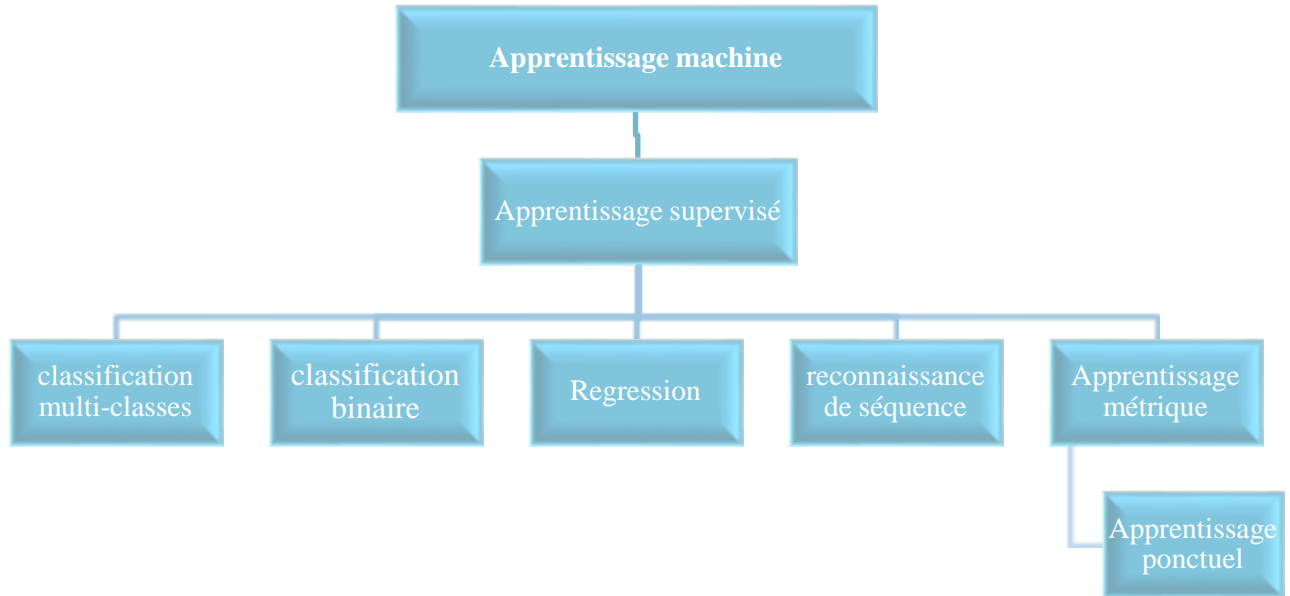


Figure 8: Les types d'apprentissage supervisé

L'apprentissage métrique est une méthode pour apprendre une représentation des données qui permet de mesurer la similarité ou la distance entre les données. Cette représentation peut ensuite être utilisée dans le cadre d'une tâche d'apprentissage supervisé, telle que la classification ou la prédiction de valeurs numériques.

Ainsi, l'apprentissage métrique et l'apprentissage supervisé sont deux approches distinctes, mais complémentaires, qui peuvent être utilisées ensemble pour améliorer les performances d'un modèle.

6.2 Apprentissage non supervisé

L'apprentissage non supervisé est une méthode de transformation de données qui ne nécessite pas de connaître à l'avance l'ensemble de données cible.

Contrairement à l'apprentissage supervisé, qui utilise des données étiquetées, l'apprentissage non supervisé se nourrit uniquement d'exemples et utilise des techniques telles que le clustering pour créer des groupes logiques ou des règles d'association telles que les algorithmes a priori.

Un exemple d'algorithme couramment utilisé pour la classification non supervisée est K-means, qui facilite la compréhension du concept de cette méthode. [28]

6.3 Apprentissage par renforcement

Dans un système d'apprentissage par renforcement, le modèle n'est pas fourni avec des exemples d'entrées et de sorties.

Au lieu de cela, il est informé de l'état actuel du système, d'un objectif à atteindre, d'une liste d'actions possibles et de leurs résultats attendus en fonction des contraintes environnementales.

Le modèle doit alors tester par lui-même le processus pour atteindre l'objectif, en utilisant le principe d'essais et d'erreurs pour maximiser une récompense. [28]

7 Apprentissage Métrique

L'apprentissage métrique est une branche de l'apprentissage automatique qui se concentre sur la création de modèles qui peuvent apprendre à mesurer les similitudes et les différences entre les données en fonction de leurs caractéristiques. Cette technique est souvent utilisée pour la classification, la reconnaissance de formes et le clustering. [29]

L'apprentissage métrique existe déjà et a été utilisé dans de nombreuses applications telles que la reconnaissance faciale, la recherche d'images et la recommandation de produits. L'apprentissage métrique profond, qui utilise des réseaux de neurones profonds pour apprendre des mesures ou des distances entre les données, est également une réalité et est utilisé dans des domaines tels que la reconnaissance de la parole, la reconnaissance d'images et la recommandation de produits. [30]

7.1 L'apprentissage ponctuel (One shot Learning)

One shot learnig (L'apprentissage ponctuel) est un sous-domaine de l'apprentissage automatique qui se concentre sur la formation de modèles pour faire des prédictions précises basées sur un seul ou très peu d'exemples par classe. Il relève le défi de l'apprentissage à partir de données limitées, où les algorithmes d'apprentissage traditionnels peuvent avoir du mal à généraliser efficacement.

L'objectif d'apprentissage ponctuel est de développer des modèles capables de reconnaître et de classer de nouvelles instances après avoir vu un seul exemple de chaque classe pendant la formation. Cette capacité à généraliser à partir de données minimales est particulièrement utile dans les scénarios où l'acquisition de données étiquetées est coûteuse ou prend du temps.

Les méthodes de L'apprentissage ponctuel s'appuient souvent sur des techniques telles que l'apprentissage métrique, les réseaux siamois ou les modèles génératifs pour capturer les caractéristiques essentielles ou les similitudes entre les exemples.

L'apprentissage ponctuel trouve des applications dans divers domaines, notamment la reconnaissance faciale, la reconnaissance d'objets, la reconnaissance de l'écriture manuscrite et les systèmes de recommandation personnalisés. Il permet le développement de modèles capables de faire des prédictions précises avec un minimum de données d'entraînement, réduisant ainsi la dépendance à l'égard de grands ensembles de données étiquetées.

- **Comparaison entre One shot learnig et classification**

La classification et One shot learnig sont deux tâches d'apprentissage automatique qui impliquent l'attribution d'étiquettes aux données. Cependant, ils diffèrent par la quantité de données de formation requises.

En classification, le modèle est formé sur un grand ensemble de données de données étiquetées. Cela permet au modèle d'apprendre les caractéristiques qui distinguent les

différentes classes. Une fois le modèle formé, il peut être utilisé pour classer de nouvelles données avec un haut degré de précision.

L'apprentissage ponctuel (One shot learning), en revanche, ne nécessite qu'un seul exemple de formation pour chaque classe. Cela rend la tâche plus difficile, car le modèle doit apprendre à généraliser à partir d'une très petite quantité de données. Cependant, l'apprentissage ponctuel peut être utile dans des situations où il est difficile ou impossible de collecter un grand ensemble de données étiquetées.

L'apprentissage ponctuel est souvent utilisé dans les tâches de vision par ordinateur, telles que la reconnaissance faciale et la détection d'objets. Dans ces tâches, il peut être difficile de collecter un grand ensemble de données de données étiquetées. L'apprentissage ponctuel peut être utilisé pour former des modèles qui peuvent encore atteindre une grande précision, même avec une quantité limitée de données de formation.

7.2 Les modèles d'apprentissage métrique

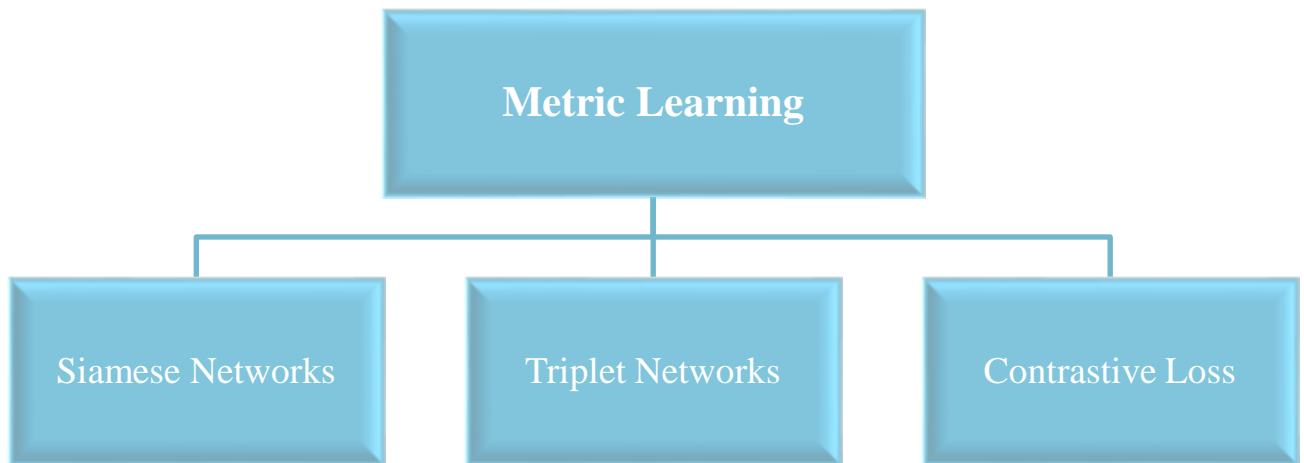


Figure 9: Les différents modèles utilisés dans Metric Learning

Il existe plusieurs modèles utilisés en Metric Learning, voici quelques exemples :

7.2.1. Siamese Networks

Il s'agit d'un réseau de neurones qui prend en entrée deux images et qui apprend à les comparer. Les deux images sont passées à travers des couches de convolution partagées avant d'être transformées en vecteurs de caractéristiques qui sont ensuite comparés. [31]

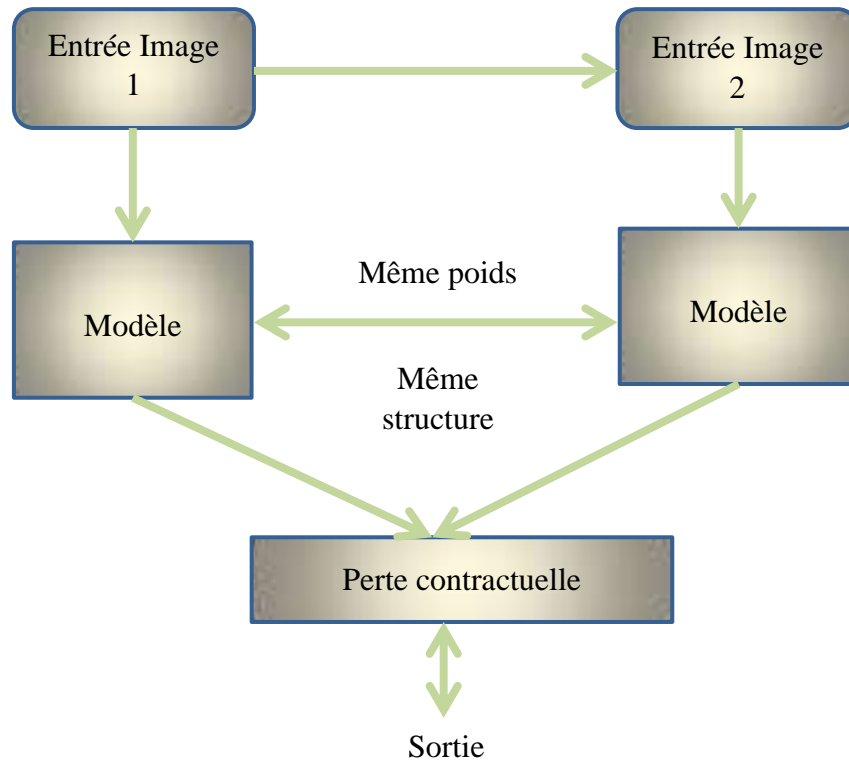


Figure 10: le modèle de réseau siamoise [32]

7.2.2. Contrastive Loss

Il s'agit d'un algorithme qui utilise des paires d'images et assigne une distance souhaitée pour des images similaires et une distance minimale pour des images différentes. L'objectif de l'apprentissage est de minimiser la distance entre les vecteurs de caractéristiques pour des images similaires et de maximiser la distance pour des images différentes. [34]

La fonction mathématique de perte (ou "loss" en anglais) pour l'apprentissage construit est souvent appelée "Contrastive Loss" en anglais. Elle est utilisée dans les tâches d'apprentissage automatique où l'on cherche à apprendre une représentation compacte d'un espace de données en maximisant la similarité entre des exemples similaires et en minimisant la similarité entre des exemples différents.

La forme générale de la fonction de perte contrastive peut varier, mais une formulation courante est la suivante :

$$L_loss(x1, x2, y) = y * d(x1, x2) + (1 - y) * \max(0, m - d(x1, x2))$$

où :

- L_loss est la fonction de perte contrastive,
- $x1$ et $x2$ sont les exemples d'entrée que l'on souhaite comparer,
- y est l'étiquette binaire qui indique si les exemples sont similaires ($y = 1$) ou différents ($y = 0$),

- $d(x_1, x_2)$ est une mesure de la similarité entre x_1 et x_2 (par exemple, la distance euclidienne),
- m est un seuil qui contrôle la marge de séparation entre les exemples similaires et différents.

La première partie de l'expression ($y * d(x_1, x_2)$) représente la similarité entre les exemples similaires (lorsque $y = 1$), où l'on cherche à minimiser la distance entre x_1 et x_2 . La deuxième partie de l'expression ($(1 - y) * \max(0, m - d(x_1, x_2))$) représente la similarité entre les exemples différents (lorsque $y = 0$), où l'on cherche à maximiser la distance entre x_1 et x_2 avec une marge minimale de m .

La fonction de perte contrastive est souvent utilisée en conjonction avec des techniques d'apprentissage profond, telles que les réseaux de neurones siamois, pour apprendre des représentations discriminantes dans des tâches telles que la reconnaissance de visages, la recherche d'images similaires.

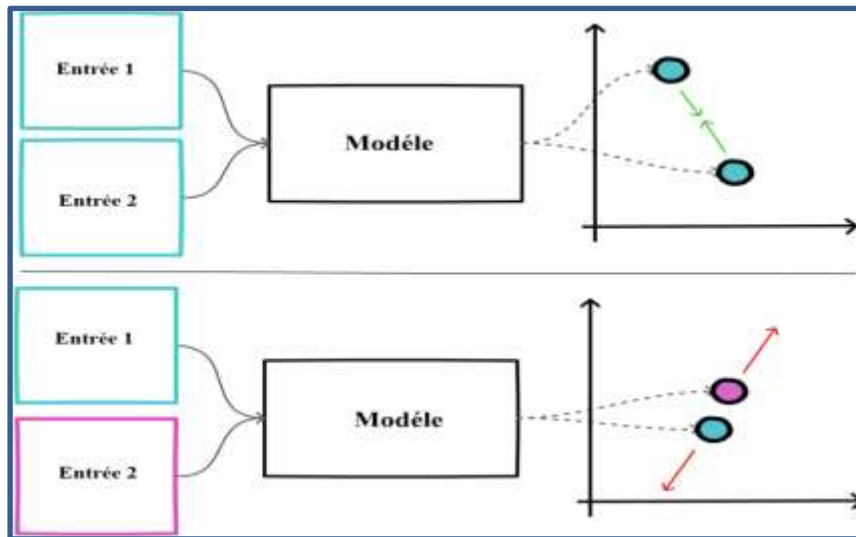
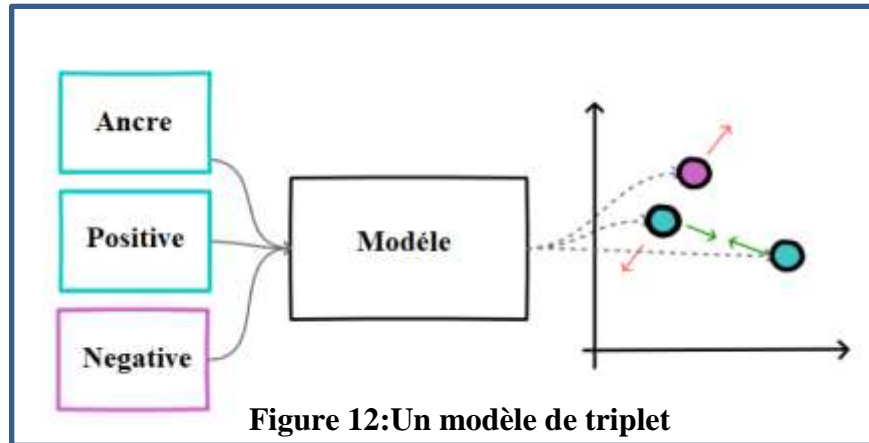


Figure 11: la perte contrastive entre les paires d'entrées

7.2.3. Triplet Networks

Il s'agit également d'un réseau de neurones qui prend trois images en entrée : une image d'ancrage, une image positive (similaire à l'image d'ancrage) et une image négative (différente de l'image d'ancrage).

L'objectif de l'apprentissage est de minimiser la distance entre les vecteurs de caractéristiques de l'image d'ancrage et de l'image positive, tout en maximisant la distance entre les vecteurs de caractéristiques de l'image d'ancrage et de l'image négative. [33]



Il existe également d'autres algorithmes utilisés en Deep Metric Learning tels que N-pair loss [34] Proxy-NCA, Angular Loss, Multi-Similarity Loss, etc.

8 IA pour les maladies des plants

8.1 Les jeux de données

Le principal défi auquel les chercheurs en apprentissage profond sont confrontés est la collecte d'ensembles de données en quantité suffisante pour obtenir la meilleure précision possible. Cette section aborde les ensembles de données accessibles au public qui contiennent des images de feuilles saines et malades de différentes plantes.

L'un des plus grands référentiels d'images de cultures en accès libre est le jeu de données PlantVillage (Hughes et Salathé, 2015). Il comprend 54 306 images de feuilles, comprenant des images de feuilles infectées et saines de 14 types de plantes différents. Toutes les images de feuilles de cet ensemble de données ont été prises avec un fond uni dans un environnement contrôlé, et les feuilles ont été étiquetées par des experts.

L'ensemble de données PlantDoc (Singh et al., 2020) se compose de 2598 images en temps réel de feuilles appartenant à 13 espèces végétales différentes, avec 27 classes, dont 17 classes pour des feuilles malades et 10 classes pour des feuilles saines. Les images de cet ensemble de données ont été capturées dans un environnement non contrôlé. Les espèces végétales incluses sont le poivron, la pomme, la myrtille, le maïs, la cerise, le raisin, la pomme de terre, la pêche, la courge, la fraise, la framboise, la tomate et le soja.

L'ensemble de données PlantDisease utilisé par Arsenovic et al. (2019) est une version étendue du jeu de données PlantVillage. Il comprend 79 265 images de 12 espèces végétales différentes, réparties en 42 classes, comprenant des classes pour les feuilles saines et malades. Cet ensemble de données comprend de nouvelles images de feuilles saines et malades capturées dans différents contextes environnementaux. Les espèces végétales présentes dans cet ensemble de données sont la tomate, la betterave à sucre, la fraise, la prune, la pomme de terre, la pêche, l'oignon, le blé, le raisin, la cerise, le poivron et la pomme.

Les ensembles de données sur les maladies de la tomate utilisés dans l'étude de Tian et al. (2019) ont été générés à l'Université agricole de Chine du Sud. Ils comprennent 1000

images de maladies foliaires, dont 200 images avec un fond blanc et 800 images avec un arrière-plan naturel. Dans cet ensemble de données, les 200 images avec un fond blanc sont exclusivement utilisées à des fins expérimentales, tandis que les 800 images restantes avec un arrière-plan réel sont utilisées à des fins de test.

Tableau 5: Les ensembles de données auto-collectés sur les maladies de plantes (tomate)

Ensemble de données	Nombre total d'images	Nombre des espèces végétales	Nombre de classes	Nombre de feuille de tomate images	Nombre de classes de tomates
Plantvillage	54,306	14	38	16529	10
PlantDoc	2,598	13	27	741	9
PlantDisease	79,265	12	42	7832	3
South China Agricultural University(Tomato)	-	-	-	1000	2

8.2 Les méthodes traditionnelles pour le traitement des maladies des plantes

Le traitement des maladies et des ravageurs des plantes à l'aide de la vision artificielle utilise généralement des algorithmes de traitement d'image conventionnels ainsi que des caractéristiques conçues manuellement et des classificateurs tels que SVM et K-Mean. [35]

Cependant, ces méthodes traditionnelles sont limitées dans leur capacité à extraire des informations complexes et détaillées à partir des images, ce qui peut rendre la reconnaissance des maladies et des ravageurs difficile et dépendante de l'expérience et de la chance.

Les algorithmes traditionnels de reconnaissance d'images sont basés sur des caractéristiques conçues manuellement, ce qui rend leur utilisation complexe et difficile.

De plus, ces algorithmes ne peuvent pas automatiquement apprendre et extraire les caractéristiques de l'image d'origine, ce qui les rend moins efficaces pour détecter les maladies et les ravageurs des plantes.

8.3 Apprentissage en profondeur pour les maladies des plantes

L'apprentissage en profondeur est une approche qui surmonte les limites des algorithmes traditionnels basés sur des fonctionnalités conçues artificiellement, et qui suscite un intérêt croissant chez les chercheurs. Cette approche a connu un grand succès dans différents domaines, tels que la vision par ordinateur, la reconnaissance de formes, la reconnaissance de la parole, le traitement du langage naturel et les systèmes de recommandation.

L'apprentissage en profondeur permet d'apprendre automatiquement des fonctionnalités à partir de grandes quantités de données, sans intervention manuelle. Les modèles d'apprentissage en profondeur sont constitués de plusieurs couches, qui offrent une grande capacité d'apprentissage autonome et d'expression de caractéristiques. Ils peuvent ainsi extraire automatiquement des caractéristiques d'images pour la classification et la reconnaissance d'images.

L'utilisation de l'apprentissage en profondeur dans la reconnaissance des maladies et des ravageurs des plantes peut jouer un rôle crucial dans l'agriculture. Actuellement, de nombreux modèles de réseaux de neurones profonds, tels que le réseau de neurones à convolution profonde (CNN), ont été développés pour cette application.

Dans l'ensemble, les tomates sont un bon choix pour une étude de cas sur l'apprentissage en profondeur pour la sécurité alimentaire, car il s'agit d'une culture vivrière majeure, elles sont sensibles à un large éventail de maladies et l'apprentissage en profondeur s'est avéré efficace pour détecter et classer les maladies de la tomate.

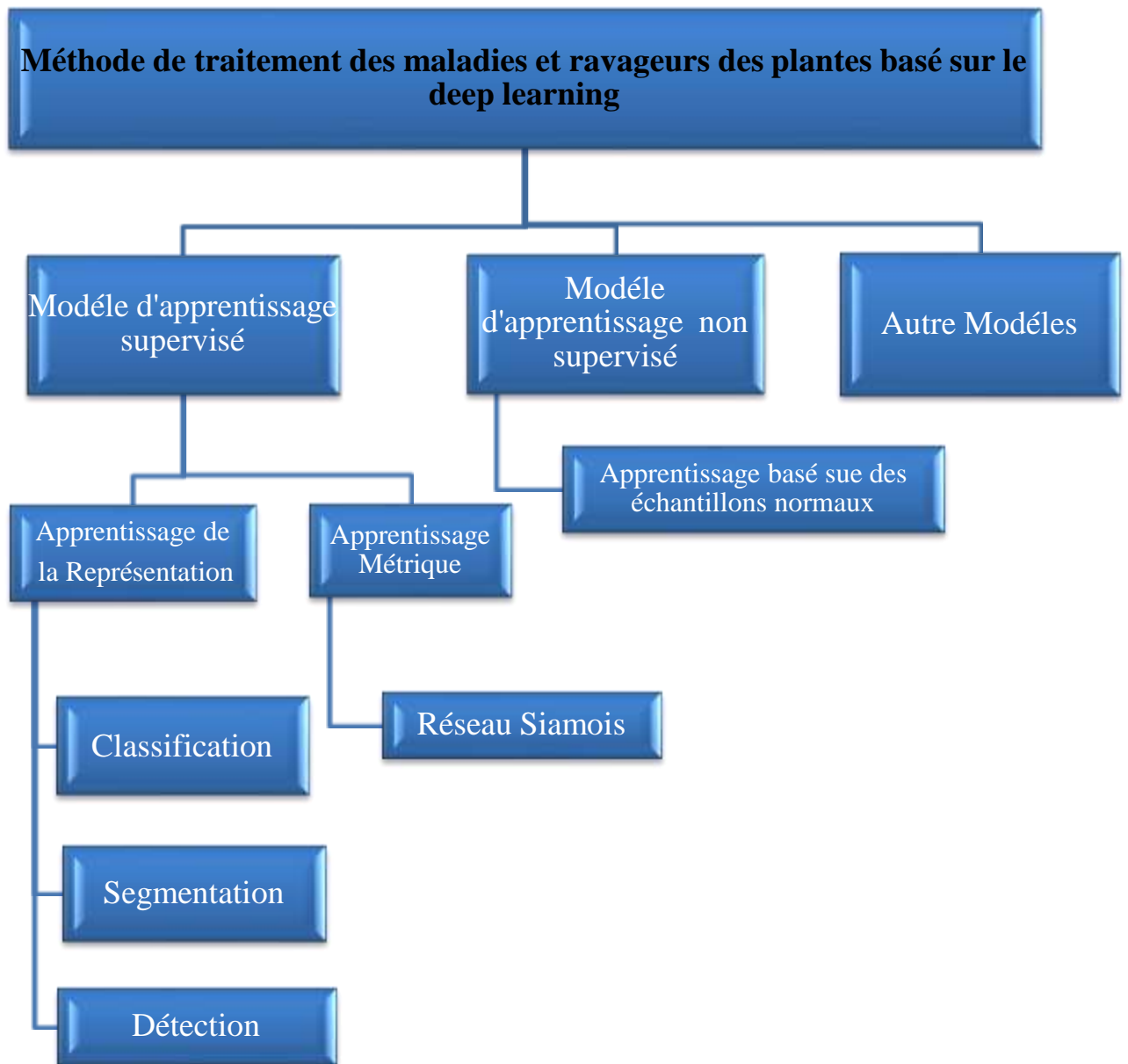


Figure 13 : Les méthodes de traitement des maladies et ravageurs de plantes basée sur le deep learning [36]

8.3.1 Apprentissage de la représentation

L'utilisation de l'apprentissage en profondeur pour traiter les maladies et les ravageurs des plantes peut être divisée en différentes structures de réseau, comme le réseau de classification, le réseau de détection et le réseau de segmentation, comme illustré dans la figure. Cette approche offre donc une alternative efficace et automatisée aux méthodes traditionnelles de reconnaissance d'images pour le traitement des maladies et des ravageurs des plantes.

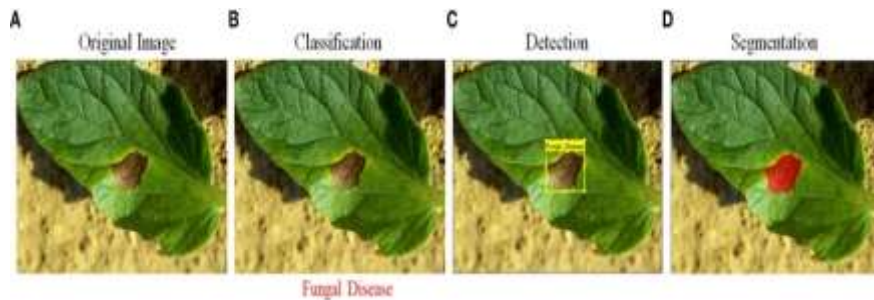


Figure 14 : Le traitement des maladies de plantes [37]

8.3.1.1 Classification

La reconnaissance des maladies et des ravageurs des plantes dans un environnement naturel est un défi en raison des variations importantes dans les formes, tailles, textures, couleurs, arrière-plans, dispositions et conditions d'éclairage des images.

Cependant, grâce à la puissance de l'extraction de caractéristiques offerte par les réseaux de neurones convolutifs (CNN), les réseaux de classification basés sur CNN sont largement utilisés pour la classification des maladies et des ravageurs des plantes. Typiquement, la partie d'extraction de caractéristiques du réseau de classification CNN est composée d'une cascade de couches de convolution et de regroupement, suivie d'une couche de connexion complète (ou de regroupement moyenne) et d'une structure softmax pour la classification.

Plusieurs structures de réseaux matures en vision par ordinateur sont utilisées pour les réseaux de classification des maladies et des ravageurs des plantes, telles que AlexNet [38], GoogleLeNet[39], VGGNet [40], ResNet [41], Inception V4 [42], DenseNets [43], MobileNet [44] et SqueezeNet [45]. Des structures de réseau spécifiques ont également été conçues pour résoudre des problèmes pratiques. Les méthodes de classification peuvent être subdivisées en trois sous-catégories :

- **Utilisation du réseau comme extracteur de fonctionnalités**

Au début des recherches sur les maladies des plantes et les méthodes de classification des ravageurs basées sur l'apprentissage en profondeur, les scientifiques ont utilisé la puissante capacité d'extraction de caractéristiques de CNN et ont combiné ces méthodes avec des classificateurs traditionnels.

Tout d'abord, les images sont introduites dans un réseau CNN pré-entraîné pour extraire les caractéristiques de l'image, qui sont ensuite utilisées comme entrées dans un classificateur d'apprentissage automatique classique (par exemple, SVM) pour effectuer la classification.

- **Utilisation directe du réseau pour la classification**

La classification directe des lésions à l'aide du réseau de classification est l'une des méthodes couramment utilisées de CNN dans la détection des maladies et des ravageurs des plantes. Cette méthode peut être divisée en trois sous-catégories en fonction des caractéristiques des travaux de recherche existants : classification d'image originale, classification après localisation de la région d'intérêt (ROI) et classification multi-catégories.

- **Utilisation du réseau pour la localisation des lésions**

En plus de la classification d'image, le réseau de classification peut également être utilisé pour réaliser la localisation des lésions et la classification pixel par pixel en combinant différentes techniques et méthodes. Selon les méthodes utilisées, il peut être subdivisé en trois formes : la méthode de la fenêtre glissante, la méthode de la carte thermique et la méthode du réseau d'apprentissage multi-tâches.



Fenêtre coulissante.

La méthode de fenêtre glissante est la plus simple et la plus intuitive pour localiser grossièrement une lésion. Elle implique de faire glisser une fenêtre de petite taille sur l'image d'origine, et d'entrer chaque portion d'image dans le réseau de classification pour détecter les maladies et les ravageurs des plantes. Les résultats de la classification pour chaque fenêtre sont finalement combinés pour obtenir les résultats de la localisation de la lésion.



Carte de chaleur.

La carte thermique est une représentation visuelle qui permet de mettre en évidence l'importance relative de chaque région d'une image. Les régions les plus importantes sont représentées en couleur foncée, tandis que les régions moins importantes sont représentées en couleur claire. Dans le domaine de la détection des maladies et des ravageurs des plantes, la couleur de la carte thermique peut être utilisée pour indiquer la probabilité qu'une région donnée de l'image soit affectée par une maladie ou un ravageur. En général, plus la couleur est foncée, plus la probabilité qu'il s'agisse de la lésion est élevée.



Réseau d'apprentissage multi-tâches.

Un réseau de classification pure ne peut localiser avec précision les maladies et ravageurs des plantes qu'au niveau de l'image. Pour remédier à cela, une branche supplémentaire peut être ajoutée au réseau, permettant de combiner la classification et la segmentation pour former un réseau d'apprentissage multitâche.

Les deux branches partagent les résultats de l'extraction de caractéristiques, réduisant ainsi les exigences en matière d'échantillonnage pour le réseau de classification. Ren et al. [46] ont construit un modèle DGVGNet basé sur VGNet pour identifier les maladies des feuilles des plantes. La déconvolution a été utilisée pour guider le classificateur CNN à se concentrer sur les sites de lésion réels et à améliorer la précision de la classification et de la segmentation des lésions. Les résultats des tests ont montré que le modèle a une précision de 99,19 % pour l'identification des classes de maladies, une précision de 94,66 % pour la segmentation des lésions et une bonne robustesse en cas d'occlusion, de basse lumière et dans d'autres environnements.

8.3.1.2 Détection

La localisation d'objet est une tâche fondamentale en vision par ordinateur, et elle est étroitement liée à la détection des maladies et ravageurs des plantes. Les méthodes de détection basées sur l'apprentissage profond ont connu un développement rapide. En général, il existe deux types de réseaux de détection de maladies et ravageurs des plantes basés sur l'apprentissage en profondeur : les réseaux en deux étapes, représentés par Faster R-CNN [47], et les réseaux en une étape, représentés par SSD [48] et YOLO [49].

La principale différence entre les deux types de réseaux est que les réseaux en deux étapes génèrent d'abord des boîtes candidates pour les lésions, puis effectuent la détection d'objet, tandis que les réseaux en une étape prédisent directement l'emplacement et la classe des lésions en utilisant les caractéristiques extraites par le réseau.

- **Détection des maladies et ravageurs des plantes basée sur un réseau à deux étapes**

Le réseau de détection en deux étapes, tel que le Faster R-CNN, procède en premier lieu à l'obtention de la carte des caractéristiques de l'image d'entrée à travers le réseau de base, avant de calculer la confiance de la boîte d'ancrage en utilisant la région proposée par RPN.

Le réseau traite ensuite la carte des caractéristiques de la zone proposée, via ROIpooling, pour affiner les résultats de détection initiaux et obtenir finalement les résultats de localisation et de classification des lésions.

En conséquence, les méthodes de détection de maladies et ravageurs des plantes basées sur le deep learning améliorent souvent les éléments clés tels que la structure du réseau de base, la carte des caractéristiques, le ratio d'ancrage, le ROIpooling et la fonction de perte, en fonction des caractéristiques de la tâche de détection en question.

En 2017, Fuentes et al. [50] a d'abord utilisé Faster R-CNN pour localiser directement les maladies et les ravageurs de la tomate, combiné à des extracteurs de caractéristiques profondes tels que VGG-Net et ResNet, la valeur mAP a atteint 85,98 % dans un ensemble de données contenant 5000 maladies et ravageurs de la tomate de 9 catégories.

- **Détection des maladies et des ravageurs des plantes sur la base d'un réseau à une étape**

L'algorithme de détection d'objets en une seule étape a simplifié la méthode de détection en éliminant l'étape de proposition de région. À la place, la tête de détection est ajoutée directement au réseau de base pour la classification et la régression, ce qui a considérablement accéléré la vitesse d'inférence du réseau. Il existe deux types de réseaux de détection à une étape : SSD et YOLO. Les deux réseaux prennent l'image entière en entrée et renvoient directement les informations de position de la boîte englobante et de la catégorie correspondante au niveau de la couche de sortie.

Le SSD utilise VGG16 comme tronc de réseau et ajoute un réseau pyramidal de caractéristiques pour obtenir des caractéristiques de différentes couches et faire des prédictions.

YOLO (You Only Look Once) traite la détection d'objets comme un problème de régression et utilise une seule architecture de réseau de neurones convolutifs pour prédire simultanément la boîte englobante et la catégorie de l'objet. Il considère l'ensemble de l'image

globale plutôt que des régions d'intérêt, ce qui permet une optimisation globale et une accélération significative de la détection, tout en maintenant une précision élevée. Ainsi, YOLO offre une solution de détection d'objet end-to-end efficace et performante.

Liu J, Wang X [51], L'algorithme Yolo V3 a été utilisé pour détecter les maladies de la tomate et les insectes ravageurs. Les résultats expérimentaux montrent que la précision de détection de l'algorithme est de 92,39 % et le temps de détection n'est que de 20,39 ms.

Dans le domaine de la détection des maladies et des ravageurs des plantes, les modèles de détection basés sur deux étapes sont souvent utilisés pour une détection précise, tandis que les modèles basés sur une étape sont privilégiés pour une vitesse de détection rapide. Toutefois, il est important de noter que la tâche du réseau de détection est différente de celle du réseau de classification.

Le réseau de détection permet de localiser les maladies et les ravageurs des plantes, mais ne peut pas identifier leur catégorie. Ainsi, le réseau de classification est nécessaire pour déterminer la classe des maladies et des ravageurs des plantes. Bien que le réseau de détection et le réseau de classification aient des rôles complémentaires dans la résolution des problèmes de maladies et de ravageurs des plantes, le réseau de détection ne peut pas remplacer le réseau de classification.

8.3.1.3 Segmentation

Le réseau de segmentation est une technique qui permet de transformer la tâche de détection des maladies et des ravageurs des plantes en une segmentation sémantique, permettant ainsi d'obtenir des informations plus détaillées sur les lésions et les zones normales. Cette méthode permet de diviser finement la zone de la lésion et de fournir des informations sur l'emplacement, la catégorie et les propriétés géométriques associées telles que la longueur, la largeur, la surface, le contour, le centre, etc.

Il existe deux types de réseaux de segmentation couramment utilisés : les réseaux entièrement convolutifs (FCN) [29][52] et les masques R-CNN [53].

- **FCN**

La segmentation sémantique des images est fondée sur le réseau de neurones à convolution entièrement convolutif (FCN), qui constitue une méthode de base pour la segmentation sémantique. Les caractéristiques de l'image d'entrée sont d'abord extraites et encodées par convolution dans FCN. Ensuite, l'image caractéristique est progressivement rétablie à la taille de l'image d'entrée en utilisant soit la déconvolution, soit l'échantillonnage. Presque tous les modèles de segmentation sémantique actuels s'appuient sur FCN.

- **Masque R-CNN**

Le masque R-CNN est largement utilisé pour la segmentation d'instance d'image. C'est une méthode d'apprentissage multitâche qui combine à la fois la détection et la segmentation dans un seul réseau. La segmentation d'instance peut être utile lorsqu'il y a plusieurs lésions du même type qui se chevauchent ou qui sont en contact, car elle peut séparer chaque lésion individuelle et permettre un décompte précis des lésions. En revanche, la segmentation sémantique traite souvent toutes les lésions du même type comme une seule entité.

Wang et al. [54] ont utilisé deux modèles différents, Faster R-CNN et Mask R-CNN, dans lesquels Faster R-CNN a été utilisé pour identifier la classe des maladies de la tomate et

Mask R-CNN a été utilisé pour détecter et segmenter l'emplacement et la forme de la zone infectée. Les résultats ont montré que le modèle proposé peut identifier rapidement et avec précision 11 classes de maladies de la tomate et diviser l'emplacement et la forme des zones infectées. Le masque R-CNN a atteint un taux de détection élevé de 99,64 % pour toutes les classes de maladies de la tomate.

La segmentation est une méthode qui offre des avantages par rapport aux approches de classification et de détection de réseaux en fournissant des informations détaillées sur les lésions. Néanmoins, comme pour le réseau de détection, la segmentation nécessite une quantité considérable de données annotées. De plus, les annotations doivent être effectuées de manière pixel par pixel, ce qui peut être un processus coûteux et chronophage.

8.3.2 Apprentissage Métrique

L'apprentissage métrique vise à résoudre les défis liés à la reconnaissance des maladies des feuilles de tomates, tels que le déséquilibre des classes et la disponibilité limitée de données d'entraînement. Il utilise une architecture légère qui permet d'obtenir des résultats précis tout en utilisant moins de ressources computationnelles.

Il existe un seul travail [56] ont utilisé une base de données d'images de feuilles de tomates provenant de la base de données PlantVillage, qui regroupe un grand nombre d'images de différentes maladies des plantes. Ils ont entraîné leur modèle en utilisant des paires d'images, où chaque paire est composée d'une image de feuille de tomate saine et d'une image de feuille de tomate malade. Le modèle utilise un réseau siamois pour extraire les caractéristiques des images et mesure la similarité entre les paires d'images.

Les résultats expérimentaux ont montré que le framework proposé par [56] atteint une précision 96.97% dans la reconnaissance des maladies des feuilles de tomates. Il a également démontré une meilleure performance par rapport à d'autres approches existantes, tout en étant plus léger en termes de ressources computationnelles nécessaires.

Tableau 6 : Approches basées sur l'apprentissage profond pour détecter et classer les feuilles des plantes de tomates

Article	Dataset	Type de traitement	Méthode	Objectif	Résultat
[46] Ren et al 2020	/	Classification	VGNet	identifier les maladies des feuilles des plantes qui étaient facilement perturbées par les ombres, les occlusions et l'intensité lumineuse	99,19 %
[50] Fuentes et al. 2017	PlantVillage	Détection	Faster R-CNN combiné à VGG-Net et ResNet	localiser directement les maladies et les ravageurs de la tomate	85,98 %
[51] Liu J, Wang X 2020.	Privé	Détection	YOLO	utilisé pour détecter les maladies de la tomate et les insectes ravageurs	92,39 %
[54] Wang et al 2019	Privé	Détection et Segmentation	Faster R-CNN Mask R-CNN	utilisé pour identifier la classe des maladies de la tomate utilisé pour détecter et segmenter l'emplacement et la forme de la zone infectée	99,64 %
[55] Karthik et al	PlantVillage	Détection	CNN	détecter le type d'infestation dans les feuilles de tomate	98 %

2020					
[56]	PlantVillage	Reconnaissance des maladies	SNN Basé CNN	Le modèle utilise un réseau siamois pour extraire les caractéristiques des images et mesure la similarité entre les paires d'images	96.97%
2022	Selvarajah Thuseethan, Palanisamy Vigneshwara, Joseph Charles and Chathrie Wimalasooriya				

9 Les défis

Problème de petite taille de jeu de données

- Les méthodes d'apprentissage en profondeur sont largement utilisées pour la détection des maladies et des ravageurs des plantes.
- Cependant, le principal obstacle à l'utilisation de ces méthodes dans ce domaine est la disponibilité limitée de données d'échantillons.
- Les ensembles de données auto-collectés sont souvent petits et deviennent beaucoup de travail pour étiqueter les données.
- Certaines maladies des plantes ont une incidence faible et coûtent cher à obtenir des images de maladies, ce qui signifie que seules quelques données d'entraînement sont disponibles.
- Le problème des petits échantillons limite donc l'application des méthodes d'apprentissage en profondeur dans la détection du mal

Amplification, synthèse et génération de données

- L'amplification des données est cruciale dans la formation de modèles d'apprentissage en profondeur pour la détection des maladies et des ravageurs des plantes.
- Une stratégie d'amplification des données optimisée peut améliorer l'efficacité de la détection des maladies et des ravageurs des plantes.
- La méthode la plus courante pour étendre les images de maladies et de ravageurs des plantes consiste à utiliser des techniques de traitement d'image telles que la mise en miroir, la rotation, le déplacement, la déformation, le filtrage, le réglage du contraste, etc. sur les échantillons d'origine de maladies et de ravageurs des plantes.

✚ Apprentissage par transfert et ajustement du modèle de réseau classique

- L'apprentissage par transfert (TL) transfère les connaissances acquises à partir de grands ensembles de données génériques vers des domaines spécialisés avec des quantités de données relativement faibles.
- L'apprentissage par transfert peut commencer par un modèle d'apprentissage par un ensemble de données connues similaires, puis l'affiner pour être appliqué à la détection localisée des maladies des plantes et des ravageurs.
- L'apprentissage par transfert peut réduire le coût de la formation du modèle et permettre au réseau neuronal de convolution de s'adapter à de petits échantillons de données.

✚ Lésions de petite taille en identification précoce

- La détection précoce et précise des maladies des plantes est cruciale pour maximiser le rendement.
- Cependant, la petite taille de l'objet lésion et le bruit de fond sur les images collectées peuvent rendre difficile la détection de petites cibles.
- Les multiples processus d'échantillonnage dans les réseaux d'extraction de caractéristiques profondes peuvent conduire à ignorer les objets à petite échelle.
- Les fonds complexes à grande échelle peuvent également augmenter le nombre de fausses détections, en particulier sur les images à faible résolution.
- En raison de la pénurie d'algorithmes existants, plusieurs stratégies telles que le mécanisme d'attention sont proposées pour améliorer les performances de la détection de petites cibles.
- L'utilisation du mécanisme d'attention permet de rationaliser les ressources allouées pour la détection des maladies et ravageurs des plantes.
- Le mécanisme d'attention trouve rapidement la région d'intérêt et ignore les informations sans importance.
- En apprenant les caractéristiques des images, le module de mécanisme d'attention peut obtenir une image saillante et isoler l'objet de l'arrière-plan, tout en supprimant le bruit de fond dans l'image.
- La fonction Softmax est utilisée pour manipuler l'image caractéristique et la combiner avec l'image caractéristique d'origine pour obtenir de nouvelles fonctionnalités de fusion à des fins de réduction du bruit.
- Dans les études futures, des mécanismes d'attention peuvent être utilisés pour sélectionner efficacement les informations et allouer davantage de ressources à la région d'intérêt pour obtenir une détection plus précise des maladies et ravageurs des plantes.
- Karthik et al. [55] ont appliqué le mécanisme d'attention sur le réseau résiduel et ont obtenu une précision globale de 98% sur l'ensemble de données plantVillage.

Identification fine

- Les caractéristiques visuelles des maladies et ravageurs des plantes de la même classe sont assez différentes.
- Les facteurs externes tels qu'un éclairage irrégulier, une occlusion dense, un tramage d'équipement flou et d'autres interférences sont à l'origine de cette différence.
- Ces facteurs entraînent des échantillons d'images différents appartenant au même type de maladies et de ravageurs.
- Ces différences forment les caractéristiques à grain fin de "différence intra-classe".
- La détection de ces maladies et ravageurs dans des scénarios complexes est une tâche difficile de reconnaissance fine.
- La présence d'un arrière-plan perturbé rend difficile la détection des maladies et des ravageurs des plantes dans le monde réel.
- Les arrière-plans peuvent être très complexes et interférer avec les objets d'intérêt.
- Un arrière-plan propre dans les images d'entraînement ne reflète pas toujours la réalité et peut conduire à des performances médiocres en situation réelle.

Performances de détection sous l'influence de l'éclairage et de l'occlusion

Problèmes d'éclairage

- La lumière naturelle change de manière très dynamique, ce qui peut provoquer une distorsion des couleurs de l'image si la plage acceptée par la caméra est dépassée.
- Les caractéristiques apparentes des maladies et des ravageurs des plantes peuvent changer considérablement en raison de la différence d'angle de vue et de distance lors de la collecte d'images.
- Ces facteurs peuvent apporter de grandes difficultés à l'algorithme de reconnaissance visuelle.

Problème d'occlusion

- La plupart des chercheurs se concentrent uniquement sur des environnements simples lorsqu'ils reconnaissent les maladies et les ravageurs des plantes.
- Ils utilisent la méthode d'interception directe de la zone d'intérêt pour les images collectées, mais ne prennent pas en compte le problème d'occlusion.
- la précision de la reconnaissance sous occlusion est faible et la praticabilité est fortement réduite.

Problème de vitesse de détection

- Les algorithmes d'apprentissage en profondeur ont des résultats supérieurs aux méthodes traditionnelles.
- La complexité de calcul est plus élevée pour les algorithmes d'apprentissage en profondeur.

- Pour garantir la précision de détection, le modèle doit apprendre pleinement les caractéristiques de l'image.
- L'augmentation de la charge de calcul peut conduire à une vitesse de détection lente.
- Les algorithmes d'apprentissage en profondeur peuvent ne pas répondre aux besoins en temps réel.

10 Synthèse

L'apprentissage métrique à l'aide d'un réseau siamois est un outil puissant pour la détection des maladies des plantes, en particulier par rapport aux méthodes traditionnelles de détection, de segmentation et de classification.

Les méthodes de détection reposent généralement sur l'identification de régions spécifiques d'une image qui correspondent à des plantes malades, ce qui peut être difficile lorsque la maladie n'est présente que dans une petite partie de l'image. De même, les méthodes de segmentation tentent de séparer les zones malades d'une image des zones saines, mais cela peut être difficile lorsque la maladie se manifeste par de subtiles différences de couleur ou de texture.

Les méthodes de classification, quant à elles, tentent de catégoriser l'image entière comme saine ou malade en fonction de critères prédéterminés. Cela peut être difficile lorsque la maladie n'est présente que dans une petite partie de l'image ou lorsqu'il existe de nombreux types de maladies différents qui doivent être identifiés. En revanche, un réseau siamois d'apprentissage métrique peut apprendre une métrique de similarité entre les images, ce qui lui permet d'identifier les maladies même lorsque l'apparence de la plante affectée n'est que légèrement différente. En comparant l'image d'une plante malade à une bibliothèque d'images de plantes saines, le réseau peut identifier les différences spécifiques qui indiquent la maladie et fournir une évaluation plus nuancée du niveau de maladie présent.

De plus, les réseaux siamois (SNN) peuvent être formés avec des données limitées, ce qui est souvent le cas dans le domaine de la détection des maladies des plantes. En tirant parti de l'apprentissage par transfert, le réseau peut être formé sur un grand ensemble de données d'images de plantes, puis affiné sur un ensemble de données plus petit d'images de maladies des plantes. Cela peut aider à surmonter le défi des données limitées et à améliorer la précision de la détection des maladies. En conclusion, alors que les méthodes de détection, de segmentation et de classification ont leur utilité dans la détection des maladies des plantes, un réseau siamois pour l'apprentissage métrique offre plusieurs avantages en termes de précision, de polyvalence et de facilité de formation. En apprenant une métrique de similarité entre les images, les réseaux siamois (SNN) peuvent identifier les maladies même lorsque l'apparence de la plante affectée n'est que légèrement différente, ce qui en fait un outil efficace pour la détection des maladies des plantes.

Les architectures siamoises légères de ces dernières années ont montré des performances immaculées avec des données petites et déséquilibrées. Par conséquent, il est nécessaire de développer un cadre à part entière en réalisant les avantages de l'architecture siamoise comme première étape pour relever ces défis. [56]

11 Conclusion

Ces dernières années, l'apprentissage métrique profond, qui fournit une meilleure solution pour les données non linéaires grâce à des fonctions d'activation, a attiré l'attention des chercheurs dans de nombreux domaines différents.

Dans ce chapitre on a présenté une vision générale sur L'IA, l'apprentissage machine, l'apprentissage profond en donnant en détail la méthode choisie dans notre travail de recherche qui est le SNN.

CHAPITRE 03

CONTRIBUTION

Et EXPERIMENTATION

1 Introduction

Dans ce chapitre, on va chapitre présente la conception de notre système, en mettant en évidence l'architecture SNN (réseau neuronal siamois) que nous avons proposée. Nous décrivons également les outils d'implémentation utilisés et la plateforme sur laquelle notre travail a été réalisé. Enfin, nous exposons les résultats que nous avons obtenus dans le cadre de notre recherche.

2 Les systèmes intelligents pour traiter les maladies des plantes

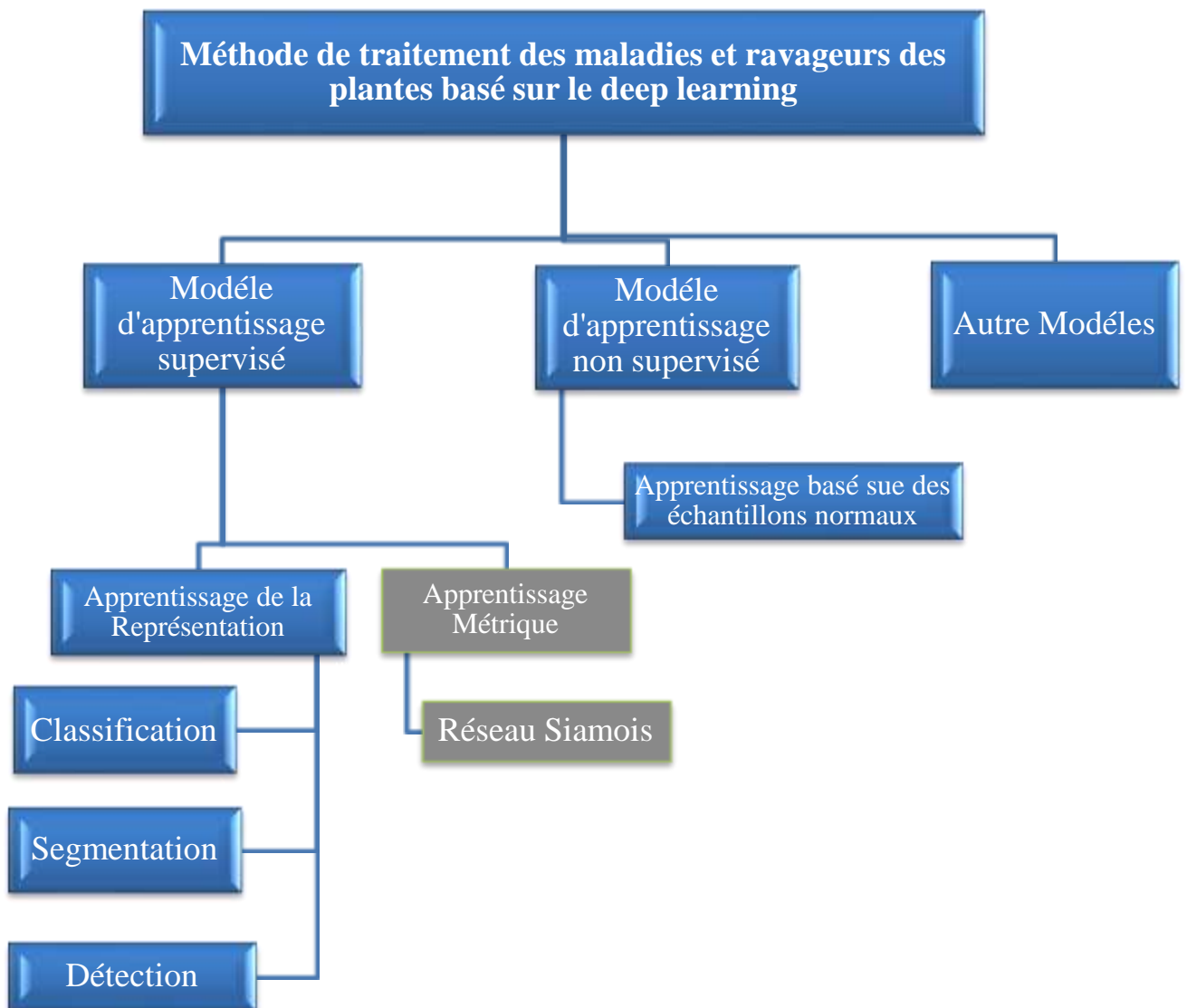


Figure 15: Les méthodes de traitement les maladies des plantes basé sur l'apprentissage métrique

Pour implémenter un modèle de réseau neuronal siamois avec l'apprentissage métrique, vous pouvez suivre les étapes suivantes :

Préparation des données : Organisez vos données en paires d'exemples, où chaque paire se compose de deux instances similaires ou non similaires. **Création du modèle siamois :** Le modèle siamois utilise deux branches identiques qui partagent les mêmes poids et architectures. Chaque branche prend en entrée une instance de la paire de données. Vous pouvez utiliser des couches convolutionnelles pour extraire des caractéristiques significatives de chaque instance. Ensuite, les caractéristiques extraites des deux branches sont combinées à l'aide d'une mesure de similarité, comme la distance euclidienne ou le produit scalaire.

Apprentissage métrique : Pour entraîner le modèle, vous devez définir une fonction de perte qui encourage la similarité entre les instances similaires et la dissimilarité entre les instances non similaires. Une perte couramment utilisée est la perte de contraste, également connue sous le nom de perte de marge. Elle calcule la distance entre les instances similaires et les instances non similaires, et vise à maximiser la marge entre ces distances.

Entraînement du modèle : Utilisez votre ensemble de données de paires pour entraîner le modèle siamois. Les instances similaires servent d'exemples positifs, tandis que les instances non similaires servent d'exemples négatifs. Appliquez la fonction de perte définie lors de l'apprentissage métrique pour guider l'apprentissage du modèle.

Évaluation du modèle : Après l'entraînement, vous pouvez évaluer le modèle en utilisant des métriques appropriées, telles que la précision.




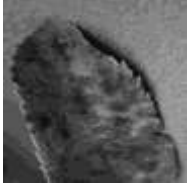
3 Jeux de données

Nous avons extrait des images de feuilles de tomates à partir de PlantVillage, une base de données récente et ouverte. Cette base de données contient un total de 54 306 images provenant de 14 cultures différentes, couvrant 38 classes de maladies pour divers types de plantes.

Notre base de données spécifique se compose de 145 29 images d'échantillons de neuf (09) classes de maladies qui peuvent affecter les plantes de tomates, ainsi qu'une classe pour les feuilles en bonne santé. Cependant, il convient de noter que notre ensemble de données est très déséquilibré en termes de distribution des échantillons d'images. La classe majoritaire dans notre ensemble de données est celle du virus de l'enroulement des feuilles jaunes, avec 4286 échantillons, soit une taille d'échantillon 14 fois plus grande que celle de la classe minoritaire, qui est la classe des maladies à virus Mosaic avec seulement 299 échantillons. Les images de l'ensemble de données d'origine ont une taille de 224×224 pixels.

Tableau 7: Description de la base de données

Désignation de classe	Nombre d'échantillons	Exemple
S 1 : Tomato Bacterial spot (Tomate Tache bactérienne)	1702	
S 2 : Tomato Early blight (Tomate Alternariose)	800	
S 3 : Tomato Late blight (Le mildiou de la tomate)	1527	
S 4 : Tomato Leaf Mold (Moisissure Feuille De Tomate)	952	
S 5 : Tomato Septoria leaf spot (Tache septorienne de la tomate)	761	
S 6 : Tomato Spider mites Two spotted (Tétranyques de la tomate Deux points)	1341	

<p>S 7 : Tomato Target Spot (Point cible de la tomate)</p>	<p>1123</p>	
<p>S 8 : Tomato yellow leaf curl virus (Virus de l'enroulement des feuilles jaunes de la tomate)</p>	<p>4286</p>	
<p>S 9 : Tomato mosaic virus (Virus de la mosaïque de la tomate)</p>	<p>299</p>	
<p>S 10 : Tomato healthy (Tomate saine)</p>	<p>1273</p>	
<p>Totale</p>	<p>14529</p>	

4 Architecture de système

L'architecture de notre système comprend les jeux de données, la phase de prétraitement des données pour préparer les images, la phase de modélisation pour construire un modèle de détection des maladies et la phase d'évaluation pour mesurer les performances du modèle sur un ensemble de test distinct

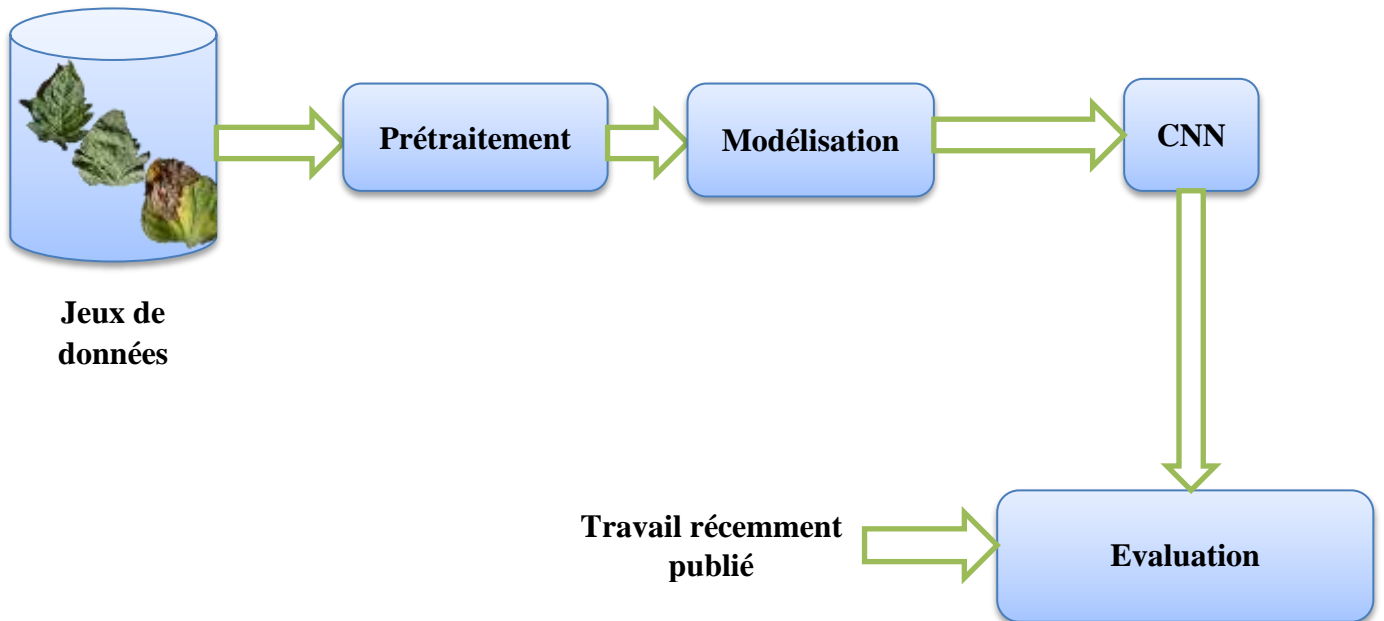


Figure 16: Conception globale

4.1 Prétraitement

Toutes les images utilisées dans la base de données étaient déjà redimensionnées à une taille fixe de 112x112 on a utilisé 50 échantillons, elles sont de niveau gris dans chaque classe. Ensuite nous avons divisé les données en 2 parties : données d'apprentissage 75% et données de test 25% du Jeu de données.

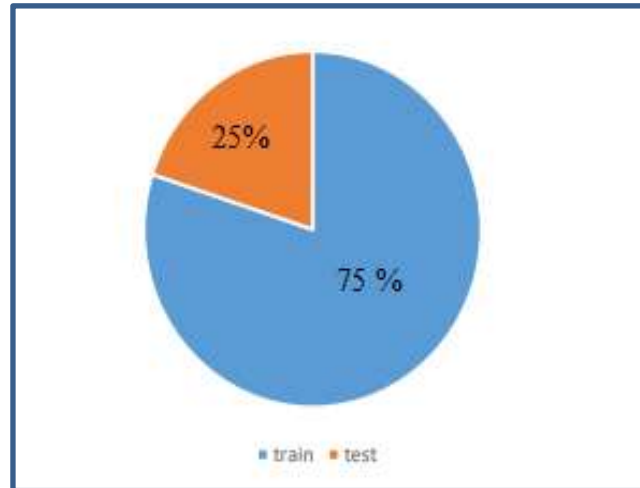


Figure 17: Division de dataset

4.2 Modélisation

Un modèle SNN (Siamese Neural Network) est un type de réseau neuronal utilisé pour mesurer la similarité ou la distance entre deux entrées. Dans ce cas, les entrées sont deux images : la première est appelée "anchor" et la deuxième est appelée "positive". L'objectif est de calculer la distance euclidienne entre les deux points correspondants dans l'espace des caractéristiques, en utilisant un CNN (Convolutional Neural Network).

Le modèle SNN suit une architecture siamoise, ce qui signifie qu'il partage les poids et les couches du CNN pour extraire les caractéristiques des deux images de manière similaire. Les images sont donc envoyées à travers le même réseau CNN pour extraire leurs caractéristiques respectives.

Une fois que les caractéristiques des deux images ont été extraites par le CNN, la distance euclidienne entre ces caractéristiques est calculée. La distance euclidienne mesure la similarité entre deux points dans un espace multidimensionnel. Elle est calculée en prenant la racine carrée de la somme des carrés des différences entre les coordonnées des points.

En utilisant la distance euclidienne, on peut évaluer à quel point les deux images sont similaires ou différentes. Une distance euclidienne plus petite indique une plus grande similarité, tandis qu'une distance euclidienne plus grande indique une plus grande dissimilarité entre les deux images. Le modèle SNN avec CNN permet donc de comparer les caractéristiques extraites des images et de quantifier leur similitude en calculant la distance euclidienne entre les points correspondants dans l'espace des caractéristiques.

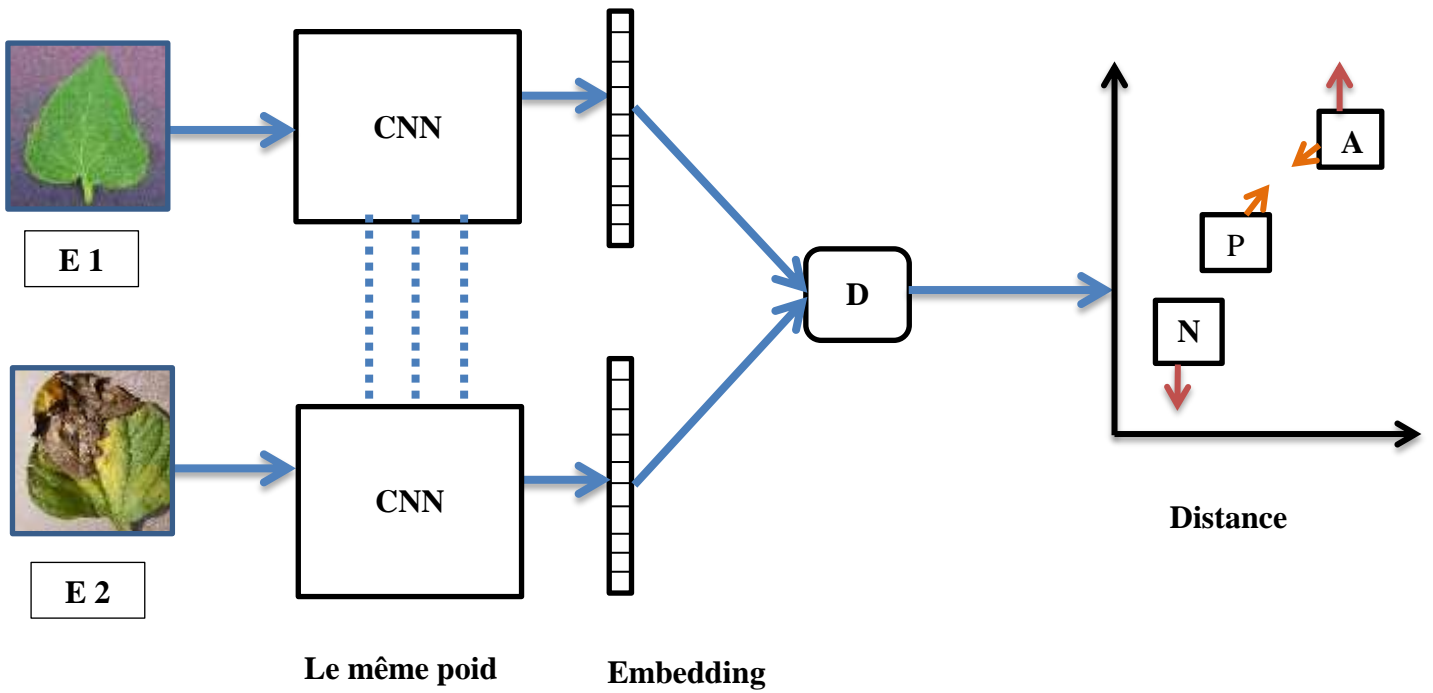


Figure 18: La figure illustre l'architecture SNN.

4.2.1 Modèle : CNN (from scratch)

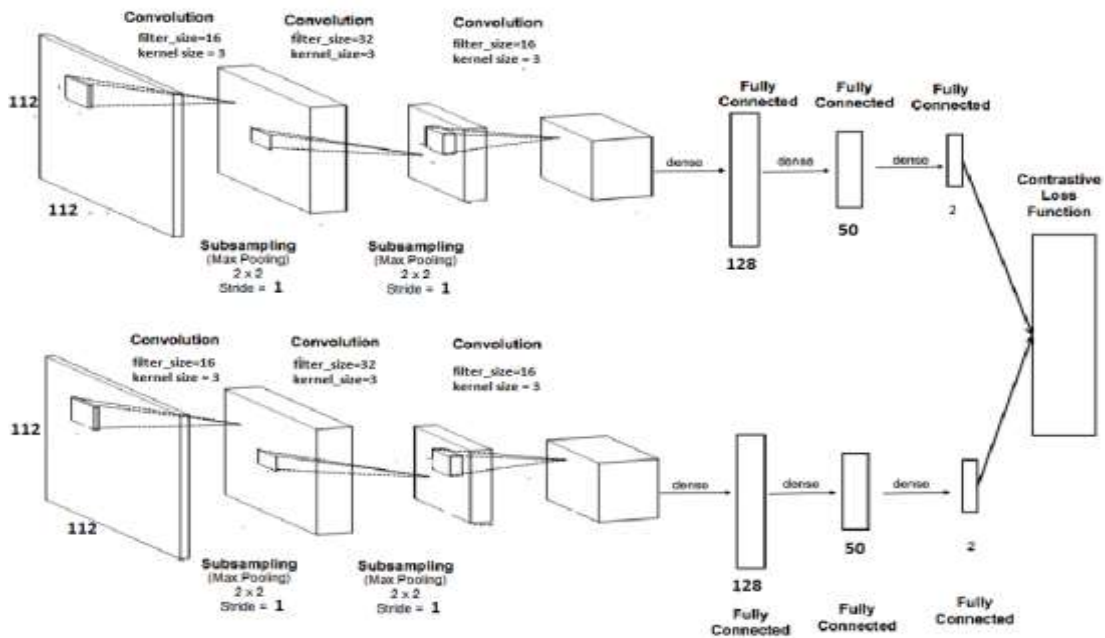


Figure 19: Le Modèle CNN

Les réseaux profonds de type CNN est largement utilisé et recommande pour le problème de classification, pour cette raison nous avons proposé un réseau convolutionnel

(CNN) comporte 12 couches : Une couche d'entrée, une couche de sortie, 3 couches de convolution, 3 couches de MaxPooling, une couche Flatten et 3 couches Fully-Connected.

- **Les couches de convolution**

Notre architecture comprend 3 couches de convolution laquelle :

- La première couche : (conv2D) est destiné à l'extraction d'identité (feature extraction).il effectue une convolution avec 16 filtres de taille 3x3 pour filtrer une image d'entrée de taille 112*112*1, la fonction d'activation utilisée est RELU.

- La deuxième couche : occupe la troisième position dans le modèle.il présente de nombreuses similitudes avec la première couche. Cette couche prend en entrée la sortie de la première couche convolutionnelle après une couche pooling avec 32 filtres de taille 3x3.

- La troisième couche occupe la cinquième position du modèle avec 16 filtres de taille 3x3 avec l'application d'une fonction d'activation RELU.

- **Les couches pooling**

Notre architecture se compose de 2 couches de pooling (maxPooling) pour réduire la dimension spatiale (feature map) des couches convolutionnelles Toutes les trois couches de pooling ont la même taille de filtre qui est 2x2.

- **Les couches fully connected (dense)**

Notre modèle se termine par une couche flatten et trois couches entièrement connectées (FC) Chaque couche utilisée une fonction d'activation RELU suivi par une couche dropout pour réduire le problème d'overfitting.

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 1, 112, 112)	0	
input_2 (InputLayer)	(None, 1, 112, 112)	0	
sequential_1 (Sequential)	(None, 50)	652674	input_1[0][0] input_2[0][0]
lambda_1 (Lambda)	(None, 1)	0	sequential_1[1][0] sequential_1[2][0]
Total params: 652,674			
Trainable params: 652,674			
Non-trainable params: 0			

Figure 20: Les couches de modèle

5 Implémentation

Nous avons utilisé diverses bibliothèques, environnements de développement et langages de programmation pour implémenter notre modèle.

5.1 Le langage de programmation :

Python est un langage de programmation multiplateforme, interprété et prenant en charge plusieurs paradigmes. Son but est de faciliter la productivité des programmeurs en proposant une syntaxe simple et des outils de haut niveau. Le langage offre un typage dynamique fort, une gestion automatique de la mémoire grâce à un ramasse-miettes et un système de gestion d'exceptions. Python est également distribué sous une licence libre et peut être utilisé sur la plupart des plateformes informatiques.

5.1.1. Les librairies et les bibliothèques Python



Figure 21: Editeur de code et bibliothèques Python utilisés

- ✚ **Scikit-learn** : Scikit-learn [58] est une bibliothèque Python qui propose une large gamme d'algorithmes d'apprentissage automatique. Elle offre des fonctions pour estimer la régression logistique, ainsi que des algorithmes de classification et de machines à vecteurs de support
- ✚ **Keras** : Keras est un framework d'apprentissage profond open source pour Python, qui peut être exécuté sur TensorFlow. L'objectif initial de Keras était de faciliter et accélérer les expérimentations avec les réseaux de neurones profonds en offrant une grande extensibilité, une compréhension facile, des performances rapides et une grande simplicité d'utilisation.[59]
- ✚ **Matplotlib** : est une bibliothèque Python dédiée à la visualisation de données en 2D. Elle permet de créer des visualisations statiques et interactives personnalisées dans différents formats. La bibliothèque offre une grande variété de possibilités de visualisation.[60]

- ✚ **NumPy** : NumPy [61] est une bibliothèque Python spécialisée dans la manipulation de tableaux multidimensionnels ou matrices. Elle propose également une vaste collection de fonctions mathématiques de haut niveau pour effectuer des opérations sur ces tableaux.
- ✚ **TensorFlow** : est une plateforme open source de bout en bout dédiée à l'apprentissage automatique (Machine Learning). [62]

Ce framework de programmation pour le calcul numérique a été développé par l'équipe Google en novembre 2015. Depuis, TensorFlow est devenu très populaire et est largement utilisé pour le Deep Learning.

Il est utilisé par de grandes entreprises telles que Intel, Twitter, Lenovo et Nvidia, et de nombreux produits de Google tels que Gmail, Google Photos et la reconnaissance vocale sont basés sur ce framework. TensorFlow est apprécié pour ses nombreux avantages, notamment :

- ✓ multi-plateformes (Linux, Mac, Windows et même Android),
- ✓ temps de compilation rapides
- ✓ exécution efficace d'opérations sur les CPU, GPU ou TPU.

5.1.2. Plateforme de développement :

L'entraînement d'un réseau de neurones nécessite une grande puissance de calcul. Dans notre méthode, nous utilisons un réseau siamois, ce qui implique un grand nombre de points à entraîner. Si nous avions utilisé un ordinateur portable pour l'entraînement, cela aurait pris beaucoup de temps. C'est pourquoi nous avons choisi d'utiliser Kaggle pour bénéficier de sa puissance de calcul et accélérer le processus d'entraînement.

- ✚ **Kaggle** : Kaggle est une communauté en ligne et une plateforme pour les passionnés de science des données et d'apprentissage automatique. Il a été fondé en 2010 et acquis par Google en 2017. Kaggle fournit une large gamme d'outils et de ressources pour les scientifiques des données, notamment des ensembles de données, des concours, des blocs-notes et un environnement collaboratif pour partager et discuter de projets liés aux données.[63]

Voici quelques fonctionnalités et activités clés sur Kaggle :

Compétitions : Kaggle organise des compétitions d'apprentissage automatique où les participants s'affrontent pour créer les meilleurs modèles prédictifs pour des problèmes spécifiques. Ces concours impliquent souvent des ensembles de données du monde réel et couvrent divers domaines tels que la santé, la finance, la vision par ordinateur et le traitement du langage naturel.

Ensembles de données : Kaggle propose un référentiel d'ensembles de données accessibles au public couvrant divers sujets. Les utilisateurs peuvent télécharger et partager des ensembles de données, explorer ceux qui existent déjà et les utiliser pour leurs propres projets et analyses.

Blocs-notes : Kaggle fournit un environnement Jupyter Notebook qui permet aux utilisateurs de créer et de partager des blocs-notes interactifs de science des données. Les blocs-notes peuvent inclure du code, des visualisations et des explications, ce qui permet aux autres de reproduire et de développer facilement le travail.

Noyaux : Kaggle Kernels est un moyen d'écrire et d'exécuter du code dans le navigateur. Les utilisateurs peuvent créer des noyaux pour partager leurs analyses, extraits de code et modèles d'apprentissage automatique. C'est un espace collaboratif où d'autres peuvent fournir des commentaires, des suggestions et des améliorations.

Discussions : Kaggle dispose d'une communauté active de scientifiques des données et de praticiens de l'apprentissage automatique. La section Discussions permet aux utilisateurs de poser des questions, de demander de l'aide et de participer à des discussions liées à la science des données et à des sujets d'apprentissage automatique.

Apprendre : Kaggle propose également des ressources pédagogiques pour l'apprentissage de la science des données et de l'apprentissage automatique. La plateforme Kaggle Learn propose des cours interactifs et des tutoriels sur divers sujets, allant des bases aux techniques avancées.

Kaggle est largement utilisé par les scientifiques des données, les chercheurs et les professionnels de l'industrie pour explorer des ensembles de données, développer des modèles d'apprentissage automatique et collaborer avec d'autres sur le terrain. C'est une plate-forme précieuse pour apprendre, pratiquer et mettre en valeur les compétences en science des données.

6 Expérimentation et Résultat

6.1 Les métriques évalué

La précision et la perte sont des mesures couramment utilisées pour évaluer les performances d'un modèle pendant les phases d'entraînement et de test.

Précision (accuracy) : la précision est une mesure de la capacité d'un modèle à prédire la sortie ou les étiquettes de classe correctes. Il est généralement exprimé en pourcentage et représente le nombre d'instances correctement prédites divisé par le nombre total d'instances. La formule de précision est :

$$\text{Précision} = (\text{Nombre d'instances correctement prédites}) / (\text{Nombre total d'instances})$$

Des valeurs de précision plus élevées indiquent de meilleures performances, une précision de 100 % représentant des prédictions parfaites.

Perte (Loss):

Dans cette méthode, le calcul de la perte contrastive L est effectué sur les paires d'échantillons d'entrée I1 et I2, contrairement à la méthode conventionnelle, où la perte est calculée comme la somme sur les échantillons d'entrée. Dans la perte contrastive, les paires d'entrée similaires sont rapprochées tandis que les échantillons d'entrée différents sont séparés. Le calcul de la perte de contraste pour la i paire d'images d'entrée est donné dans l'équation

$$L(Y,[I1,I2]^i)=(1-Y)Ls(D^i)+(Y)Ld(D^i)$$

où $L(Y,[I1,I2]^i)$ est la perte pour la i paire d'images [I1,I2] avec l'étiquette Y. La perte partielle pour des paires similaires Ls et les paires dissemblables Ld sont conçues de telle sorte que la perte globale L soit minimisée. En conséquence, le D récupère une petite valeur pour les paires d'entrées similaires et une grande valeur pour les paires d'entrées différentes. La perte contrastive résultante est donnée ci-dessous

$$L= (1-Y)1/2(D)^2+(Y)1/2\{\max(0,m-D)\}^2$$

ici, Y indique une étiquette binaire donnée à une paire d'images d'entrée [I1, I2]. Pour les paires d'entrées similaires et dissemblables, Y a les valeurs 0 et 1, respectivement. m est la marge fixée à une valeur supérieure à 0. Pour les paires dissemblables, toute distance D dans la marge m (c'est-à-dire $m - D > 0$) est pénalisée. Dans ce cas, la marge m est fixée à 2 définit la tendance à la perte des paires dissemblables. De même, les paires similaires contribuent à la perte contrastive si la distance $D > 0$.

6.2 Résultat

Dans cette section, je présente les résultats obtenus lors de la phase d'apprentissage du modèle proposé, après sa compilation. Le premier graphique représente les précisions de 75% de données d'entraînement et de 25% validation en fonction du nombre d'époques. On peut observer que la précision atteint un niveau élevé de 99,73% après 20 époques.

Le deuxième graphique illustre les pourcentages d'erreur (Loss) de 75% de données d'entraînement et de 25% validation. On remarque que ces pourcentages diminuent rapidement à mesure que la précision du modèle augmente, ce qui est tout à fait normal puisque le modèle est en phase d'entraînement. Lorsque la précision s'améliore, l'erreur de classification diminue rapidement.

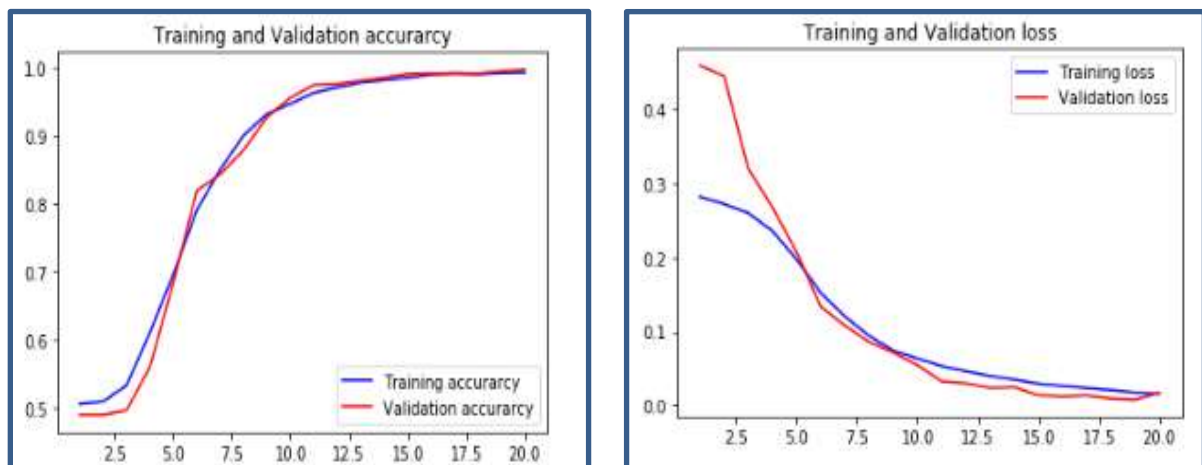


Figure 22: graphe de précision et de perte

Une fois que notre modèle est validé et prêt à être exécuté nous pouvons recevoir deux images en entrée et obtenir en sortie du système un rapport de prédiction à l'aide du modèle entraîné sur une paire d'images de test spécifique, puis compare la valeur de prédiction avec un seuil (0.5) pour décider si les images sont similaires ou non. Ensuite, il affiche à la fois la valeur de vérité terrain et la prédiction du modèle pour évaluer la performance de la prédiction.

6.3 TEST

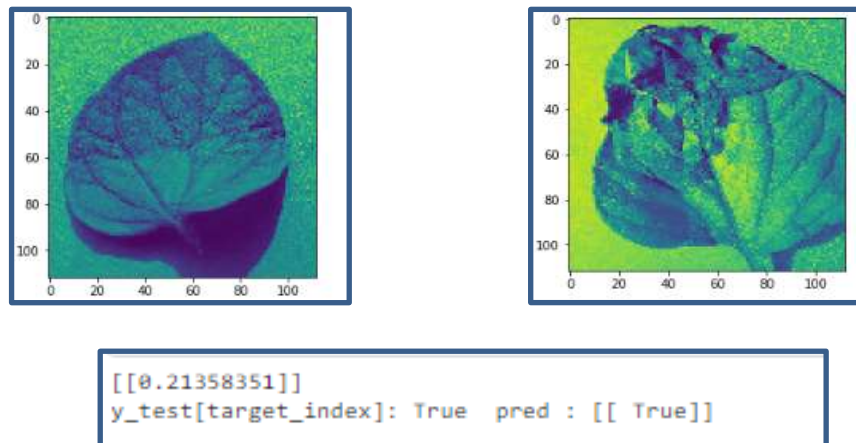


Figure 23: Etats de sortie de système

Le résultat [0.21358351] correspond à la valeur de prédiction renvoyée par le modèle. Cette valeur représente la probabilité que les deux images de test soient similaires, où une valeur plus proche de 1 indique une plus grande probabilité de similarité.

Ensuite, `y_test[target_index]: True` indique la valeur de vérité terrain pour cet exemple spécifique, c'est-à-dire que la vérité terrain indique que les deux images de test sont similaires.

Finalement, `pred : [[True]]` indique la prédiction du modèle après avoir appliqué un seuil de 0.5. Dans ce cas, la valeur de prédiction est supérieure à 0.5, ce qui correspond à une prédiction de similarité entre les deux images. Par conséquent, le modèle prédit correctement la similarité entre les images, comme indiqué par la vérité terrain.

En résumé, le modèle prédit avec succès la similarité entre les deux images de test, car la valeur de prédiction est supérieure au seuil (0.5) et correspond à la vérité terrain (True) indiquant que les images sont similaires.

6.4 Discussion

Le tableau ci-dessous présente une étude comparative entre les résultats obtenu grâce à notre travail réalisé et ceux présentés dans l'étude de référence [56].

Tableau 8: Etude comparative

Article	Modèle	Epochs	Nbrs des maladies	Précision Validation
[56]	CNN	10	10	96.97%
Notre travail	CNN	20	10	99.73 %

Nos résultats se sont avérés supérieurs (en termes de précision et de temps de réponse) par rapport à ceux de l'étude précédente [56].

7 Conclusion

Dans ce chapitre, nous avons décrit l'architecture globale de notre modèle. Nous avons commencé par présenter en détail notre contribution dans le prétraitement des données, la création du modèle et l'évaluation. Ensuite, nous avons fourni des informations détaillées sur les différentes couches de notre modèle CNN que nous avons développé. Nous avons également mentionné tous les outils utilisés lors de l'implémentation de notre modèle. Enfin, nous avons présenté les résultats obtenus et les avons discutés en détail.

Conclusion générale Et perspective

En conclusion, l'architecture SNN basée sur un modèle CNN pour la similarité des images de feuilles de tomate présente un potentiel prometteur dans la détection des maladies des plantes. Elle permet d'identifier avec précision les maladies même lorsque l'apparence de la plante affectée présente des différences subtiles.

Grâce à l'utilisation d'un modèle CNN, l'architecture SNN peut capturer efficacement les caractéristiques spécifiques des feuilles de tomate, facilitant ainsi la comparaison et la similarité entre les images. Cette robustesse face aux variations mineures dans l'apparence des feuilles est cruciale pour détecter précocement les maladies qui peuvent présenter des symptômes subtils au début.

La capacité du modèle à identifier les maladies des plantes de tomates, même avec de légères différences d'apparence, en fait un outil précieux pour les agriculteurs et les chercheurs. Une détection précoce et précise des maladies permet de prendre rapidement des mesures pour prévenir la propagation des infections, minimiser les pertes de récolte et optimiser la gestion des cultures.

Il est important de souligner que cette architecture SNN basée sur un modèle CNN peut être étendue à d'autres plantes, offrant ainsi une solution polyvalente pour la détection des maladies végétales. Cette approche peut contribuer à surveiller et à protéger différentes cultures, améliorant ainsi l'efficacité et la durabilité de l'agriculture.

Perspective

- l'amélioration des données d'entraînement.
- l'extension à d'autres plantes.
- l'application sur le terrain et la collaboration avec des experts agricoles.

BIBLIOGRAPHIE

- [1] « wasted-food-IP.pdf ». Consulté le: 9 mai 2023. [En ligne]. Disponible sur: <https://www.cccrc.org/wp-content/uploads/sites/24/2014/03/wasted-food-IP.pdf>
- [2] F. G. F. Losses et F. Waste, « Extent, causes and prevention », *Rome: Food and Agriculture Organization of the United Nations*, 2011.
- [3] « Global food security – Issues, challenges and technological solutions », *Trends in Food Science & Technology*, vol. 77, p. 11-20, juill. 2018, doi: 10.1016/j.tifs.2018.05.002.
- [4] S. McGuire, « FAO, IFAD, and WFP. The State of Food Insecurity in the World 2015: Meeting the 2015 International Hunger Targets: Taking Stock of Uneven Progress. Rome: FAO, 2015 », *Adv Nutr*, vol. 6, n° 5, p. 623-624, sept. 2015, doi: 10.3945/an.115.009936.
- [5] « mj760f.pdf ». Consulté le: 9 mai 2023. [En ligne]. Disponible sur: <https://www.fao.org/3/mj760f/mj760f.pdf>
- [6] « Linking-Humanitarian-And-Development-Interventions-into-A-Joint-Resilience-Continuum-Exploring-World-Food-Programme-WFP-and-International-Fund-for-Agricultural-Development-IFAD-Collaboration-on-Buildi.pdf ». Consulté le: 9 mai 2023. [En ligne]. pdf
- [7] « climate-foodsecurity.pdf ». Consulté le: 9 mai 2023. [En ligne]. Disponible sur: <http://re.indiaenvironmentportal.org.in/files/climate-foodsecurity.pdf>
- [8] « Nations Unies : Rapport de la Conférence mondiale sur l'alimentation :... - Google Scholar ». https://scholar.google.com/scholar_lookup?title=Report%20of%20the%20world%20food%20conference%2C%20Rome%2C%205%20E%208%9316%20November%201974&publication_year=1975 (consulté le 9 mai 2023).
- [9] A. Sen, « Ingredients of Famine Analysis: Availability and Entitlements* », *The Quarterly Journal of Economics*, vol. 96, n° 3, p. 433-464, août 1981, doi: 10.2307/1882681.
- [10] U. N. Food et A. Organization, « The State of Food Insecurity in the World 2001. Rome ». 2012.
- [11] « DROFEno8.pdf ». Consulté le: 10 mai 2023. [En ligne]. Disponible sur: <https://ofe.umontreal.ca/fileadmin/ofe/documents/Drofe/DROFEno8.pdf>
- [12] « Inwent_(2006)_Assurer_la_Sécurité_Alimentaire_et_Nutritionnelle.pdf ». Consulté le: 10 mai 2023. [En ligne]. Disponible sur: [https://wocatpedia.net/images/2/20/Inwent_\(2006\)_Assurer_la_S%C3%A9curit%C3%A9_Alimentaire_et_Nutritionnelle.pdf](https://wocatpedia.net/images/2/20/Inwent_(2006)_Assurer_la_S%C3%A9curit%C3%A9_Alimentaire_et_Nutritionnelle.pdf)
- [13] S. Sood et H. Singh, « Computer Vision and Machine Learning based approaches for Food Security: A Review », *Multimed Tools Appl*, vol. 80, n° 18, p. 27973-27999, juill. 2021, doi: 10.1007/s11042-021-11036-2.
- [14] J. Liu et X. Wang, « Plant diseases and pests detection based on deep learning: a review », *Plant Methods*, vol. 17, n° 1, p. 22, févr. 2021, doi: 10.1186/s13007-021-00722-9.

- [15] « Les principales maladies des plantes - Fiche Jardinage », *Greenastic*. <https://greenastic.com/fiche/principales-maladies-plantes/> (consulté le 10 mai 2023).
- [16] S. Khan et M. Narvekar, « Disorder Detection in Tomato Plant Using Deep Learning », in *Advanced Computing Technologies and Applications*, H. Vasudevan, A. Michalas, N. Shekokar, et M. Narvekar, Éd., in *Algorithms for Intelligent Systems*. Singapore: Springer, 2020, p. 187-197. doi: 10.1007/978-981-15-3242-9_19.
- [17] S. Khan, M. Narvekar, A. A. Shaikh, H. Ansari, et N. Ansari, « Disorder Detection in Tomato Plant Using Deep Learning ». Rochester, NY, 24 février 2019. doi: 10.2139/ssrn.3358226.
- [18] G. B. Lucas, C. L. Campbell, et L. T. Lucas, « Causes of Plant Diseases », in *Introduction to Plant Diseases: Identification and Management*, G. B. Lucas, C. L. Campbell, et L. T. Lucas, Éd., Boston, MA: Springer US, 1992, p. 9-14. doi: 10.1007/978-1-4615-7294-7_2.
- [19] A. Devaraj, K. Rathan, S. Jaahnavi, et K. Indira, « Identification of Plant Disease using Image Processing Technique », in *2019 International Conference on Communication and Signal Processing (ICCSP)*, avr. 2019, p. 0749-0753. doi: 10.1109/ICCSP.2019.8698056.
- [20] J. Karthika, K. M. kumar, M. Santhosé, T. Sharan, et S. S. hariharan, « Disease Detection In Cotton Leaf Spot Using Image Processing », *J. Phys.: Conf. Ser.*, vol. 1916, n° 1, p. 012224, mai 2021, doi: 10.1088/1742-6596/1916/1/012224.
- [21] I. Karouche et S. Bouhenni, « Classification algorithms for plant disease forecasting », *Article researchgate*, 2017.
- [22] S. Sankaran, A. Mishra, R. Ehsani, et C. Davis, « A review of advanced techniques for detecting plant diseases », *Computers and Electronics in Agriculture*, vol. 72, n° 1, p. 1-13, juin 2010, doi: 10.1016/j.compag.2010.02.007.
- [23] « Intelligence artificielle (IA) », *Data Analytics Post*. <https://dataanalyticspost.com/Lexique/intelligence-artificielle-ia/> (consulté le 10 mai 2023).
- [24] <https://www.facebook.com/atul.harsha>, « AI vs Machine Learning vs Deep Learning », *Edureka*, 8 juin 2018. <https://www.edureka.co/blog/ai-vs-machine-learning-vs-deep-learning/> (consulté le 11 mai 2023).
- [25] L. Volant, « 1996 LA PREMIÈRE RECONNAISSANCE ».
- [26] Z. ISMAILI, « Apprentissage Supervisé Vs. Non Supervisé », *BrightCape*, 28 janvier 2019. <https://brightcape.co/apprentissage-supervise-vs-non-supervise/> (consulté le 10 mai 2023).
- [27] L. Deng et D. Yu, « Deep Learning: Methods and Applications », *SIG*, vol. 7, n° 3-4, p. 197-387, juin 2014, doi: 10.1561/20000000039.

- [28] « 2018_ChaireAgroTIC_DeepLearning_VD2.pdf ». Consulté le: 10 mai 2023. [En ligne]. Disponible sur: https://www.agrotic.org/wp-content/uploads/2018/12/2018_ChaireAgroTIC_DeepLearning_VD2.pdf
- [29] R. Richard, « Apprentissage de métriques », *24pm Academy*, 4 octobre 2019. <https://24pm.com/117-definitions/497-apprentissage-de-metriques> (consulté le 11 mai 2023).
- [30] M. Kaya et H. Ş. Bilge, « Deep Metric Learning: A Survey », *Symmetry*, vol. 11, n° 9, Art. n° 9, sept. 2019, doi: 10.3390/sym11091066.
- [31] S. Benhur, « A friendly Introduction to Siamese Networks », *Medium*, 5 juin 2022. <https://towardsdatascience.com/a-friendly-introduction-to-siamese-networks-85ab17522942> (consulté le 11 mai 2023).
- [32] Prabhnoor, "Siamese Network Keras," *Medium*, [En ligne]. Disponible sur: <https://medium.com/@prabhnoor0212/siamese-network-keras-31a3a8f37d04>.
- [33] R. Gandhi, « Siamese Network & Triplet Loss », *Medium*, 17 avril 2020. <https://towardsdatascience.com/siamese-network-triplet-loss-b4ca82c1aec8> (consulté le 11 mai 2023).
- [34] S.-H. Tsang, « Review: Improved Deep Metric Learning with Multi-class N-pair Loss Objective (N-pair-mc Loss) », *Medium*, 16 novembre 2021. <https://sh-tsang.medium.com/review-improved-deep-metric-learning-with-multi-class-n-pair-loss-objective-n-pair-mc-loss-2463a7514d8> (consulté le 11 mai 2023).
- [35] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, et R. Pramodhini, « Plant leaf disease detection using computer vision and machine learning algorithms », *Global Transitions Proceedings*, vol. 3, n° 1, p. 305-310, juin 2022, doi: 10.1016/j.gltip.2022.03.016.
- [36] J. Liu et X. Wang, « Plant diseases and pests detection based on deep learning: a review », *Plant Methods*, vol. 17, n° 1, p. 22, févr. 2021, doi: 10.1186/s13007-021-00722-9.
- [37] M. Shoaib *et al.*, « An advanced deep learning models-based plant disease detection: A review of recent research », *Frontiers in Plant Science*, vol. 14, 2023, Consulté le: 11 mai 2023. [En ligne]. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1158933>
- [38] M. Z. Alom *et al.*, « The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches ». arXiv, 12 septembre 2018. doi: 10.48550/arXiv.1803.01164.
- [39] F. Yuesheng *et al.*, « Circular Fruit and Vegetable Classification Based on Optimized GoogLeNet », *IEEE Access*, vol. 9, p. 113599-113611, 2021, doi: 10.1109/ACCESS.2021.3105112.
- [40] U. Muhammad, W. Wang, S. P. Chattha, et S. Ali, « Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification », in *2018 24th International Conference*

- on *Pattern Recognition (ICPR)*, août 2018, p. 1622-1627. doi: 10.1109/ICPR.2018.8545591.
- [41] S. Targ, D. Almeida, et K. Lyman, « Resnet in Resnet: Generalizing Residual Architectures ». arXiv, 25 mars 2016. doi: 10.48550/arXiv.1603.08029.
- [42] C. Szegedy, S. Ioffe, V. Vanhoucke, et A. Alemi, « Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning », *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, n° 1, Art. n° 1, févr. 2017, doi: 10.1609/aaai.v31i1.11231.
- [43] K. Zhang, Y. Guo, X. Wang, J. Yuan, et Q. Ding, « Multiple Feature Reweight DenseNet for Image Classification », *IEEE Access*, vol. 7, p. 9872-9880, 2019, doi: 10.1109/ACCESS.2018.2890127.
- [44] L. G. Falconí, M. Pérez, et W. G. Aguilar, « Transfer Learning in Breast Mammogram Abnormalities Classification With Mobilenet and Nasnet », in *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, juin 2019, p. 109-114. doi: 10.1109/IWSSIP.2019.8787295.
- [45] B. Koonce, « SqueezeNet », in *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, B. Koonce, Éd., Berkeley, CA: Apress, 2021, p. 73-85. doi: 10.1007/978-1-4842-6168-2_7.
- [46] R. Shougang, J. Fuwei, G. Xingjian, Y. Peishen, X. Wei, "Deconvolution-guided tomato leaf disease identification and lesion segmentation model," *Journal of Agricultural Engineering*, 2020.
- [47] S. Ren, K. He, R. Girshick, et J. Sun, « Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, n° 6, p. 1137-1149, juin 2017, doi: 10.1109/TPAMI.2016.2577031.
- [48] W. Liu *et al.*, « SSD: Single Shot MultiBox Detector », 2016, p. 21-37. doi: 10.1007/978-3-319-46448-0_2.
- [49] P. Jiang, D. Ergu, F. Liu, Y. Cai, et B. Ma, « A Review of Yolo Algorithm Developments », *Procedia Computer Science*, vol. 199, p. 1066-1073, janv. 2022, doi: 10.1016/j.procs.2022.01.135.
- [50] A. Fuentes, S. Yoon, S. C. Kim, et D. S. Park, « A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition », *Sensors*, vol. 17, n° 9, Art. n° 9, sept. 2017, doi: 10.3390/s17092022.
- [51] J. Liu et X. Wang, « Tomato Diseases and Pests Detection Based on Improved Yolo V3 Convolutional Neural Network », *Frontiers in Plant Science*, vol. 11, 2020, Consulté le: 11 mai 2023. [En ligne]. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00898>
- [52] E. Shelhamer, J. Long, et T. Darrell, « Fully Convolutional Networks for Semantic Segmentation », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, n° 4, p. 640-651, avr. 2017, doi: 10.1109/tpami.2016.2572683.

- [53] K. He, G. Gkioxari, P. Dollar, et R. Girshick, « Mask R-CNN », *2017 IEEE International Conference on Computer Vision (ICCV)*, p. 2980-2988, oct. 2017, doi: 10.1109/ICCV.2017.322.
- [54] Q. Wang, F. Qi, M. Sun, J. Qu, et J. Xue, « Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and Object Detection Techniques », *Computational Intelligence and Neuroscience*, vol. 2019, p. e9142753, déc. 2019, doi: 10.1155/2019/9142753.
- [55] K. R., H. M., S. Anand, P. Mathikshara, A. Johnson, et M. R., « Attention embedded residual CNN for disease detection in tomato leaves », *Applied Soft Computing*, vol. 86, p. 105933, janv. 2020, doi: 10.1016/j.asoc.2019.105933.
- [56] S. Thuseethan, P. Vigneshwaran, J. Charles, et C. Wimalasooriya, « Siamese Network-based Lightweight Framework for Tomato Leaf Disease Recognition ». arXiv, 18 septembre 2022. Consulté le: 27 mai 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/2209.11214>
- [57] A. Kamilaris et F. X. Prenafeta-Boldú, « Deep learning in agriculture: A survey », *Computers and Electronics in Agriculture*, vol. 147, p. 70-90, avr. 2018, doi: 10.1016/j.compag.2018.02.016.
- [58] « scikit-learn Tutorials », *scikit-learn*. <https://scikit-learn/stable/tutorial/index.html> (consulté le 13 mai 2023).
- [59] « Home - Keras Documentation ». <https://faroit.com/keras-docs/1.2.0/> (consulté le 13 mai 2023).
- [60] « introduction-aux-graphiques-en-python-avec-matplotlib-pyplot.pdf ». Consulté le: 13 mai 2023. [En ligne]. Disponible sur: <https://zestedesavoir.com/tutoriels/pdf/469/introduction-aux-graphiques-en-python-avec-matplotlib-pyplot.pdf>
- [61] « numpy-user.pdf ». Consulté le: 13 mai 2023. [En ligne]. Disponible sur: <https://numpy.org/doc/1.21/numpy-user.pdf>
- [62] « tensorflow_tutorial.pdf ». Consulté le: 13 mai 2023. [En ligne]. Disponible sur: https://www.tutorialspoint.com/tensorflow/tensorflow_tutorial.pdf
- [63] « Kaggle: Your Home for Data Science ». <https://www.kaggle.com/> (consulté le 28 mai 2023).