

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DE 8 MAI 1945 - GUELMA -
FACULTÉ DES MATHÉMATIQUES, D'INFORMATIQUE ET DES SCIENCES DE LA MATIÈRE

Département d'Informatique



Mémoire de Fin d'études Master

Filière : Informatique

Option : Système Informatique

Thème _____

Détection des maladies des pommiers sur les branches

Présenté par : TOUAHRI Mohamed anis

Membres de jury :

N	Nom	Qualité
1	Chemsenahar BENCHERIET	Président
2	Brahim FAROU	Encadrant
3	Abdelaziz BENKIRAT	Représentant de l'incubateur
4	Chokri FERKOUS	Examineur

Juin 2023

Remerciements

Je tenais à remercier tous d'abord DIEU, le tous puissant de nous avoir accordé la volonté, force, santé, détermination et courage afin d'accomplir ce modeste travail.

J'exprime notre profonde et sincère expression de remerciements a mon encadreur, Professeur FAROU brahim, d'avoir dirigé ce travail et patienter avec moi et pour la confiance qu'il m'a accordé, pour sa disponibilité, ses conseils et ses idées innovantes tout au long de notre travail. J'ai appris avec lui comment développer nos sens d'écoute, d'analyse et de communication. A travers sa longue expérience combinant l'ingénierie, la recherche scientifique, l'enseignement et finalement la direction.

Je remercie vivement les membres du jury d'avoir accepté de siéger et de juger notre travail. Nous les remercions aussi pour l'honneur qu'ils nous font en acceptant d'examiner ce mémoire.

Mes remerciements vont également à tous les enseignants de l'université et également tout le personnel du département d'informatique pour leurs gentilleses.

Enfin mes remerciements les plus sincères à ma famille et à tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste travail.

DÉDICACES

Je dédie ce modeste travail à :

Ma chère mère et mon trésor de bonheur Achouri Nadia, Mon cher père et notre fierté Touahri Fouad, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études.

Mon cher frère et mon ami Daya pour son encouragement permanent, et leur soutien moral.

Mon cher frère Chouayeb et le petit de la famille Iyed, pour leur présence dans ma vie.

A toute la famille Touahri et la famille Achouri.

A ma chère Lina. Merci d'être toujours à mes côtés dans les bons et les mauvais moments, sans compter.

à mon cher encadrant monsieur Faro brahim, pour sa bonté avec moi. Il était comme mon deuxième père. Merci pour votre engagement et votre dévouement ont contribué à rendre cette expérience éducative exceptionnelle.

à mes amis les plus dévoués Hazem bensaleh, Nadhir Oumeddour, Borhan bou-nab, Akrem boughazi, Lekehal taher, Reggame moncef, Nadhir cheleghmia, Saidi ammar, Je veux vous remercier pour votre appui et vos encouragements.

à mes amis, yacinne, moncef, hani, wassim, houssem, mohammed, seif, abdallah, amine, khalil, haitem.

ملخص

تركز رسالة الماستر هذه على كشف الأمراض على فروع أشجار التفاح باستخدام الذكاء الاصطناعي. هدف هذه الدراسة هو تطوير نظام ذكي قادر على كشف وتقسيم الأمراض تلقائيًا على فروع أشجار التفاح من خلال تحليل الصور. تعتبر طرق الكشف اليدوي التقليدية مستغرقة ومكلفة وتعاني من أخطاء بشرية. من خلال استغلال تطورات التعلم الآلي ومعالجة الصور وخوارزميات التعلم العميق، يمكن لنظام قائم على الذكاء الاصطناعي توفير معلومات دقيقة وفي الوقت الحقيقي حول حالة صحة أشجار التفاح، مما يتيح للمزارعين اتخاذ تدابير وقائية وتحسين إنتاجهم الزراعي. يجمع النظام المقترح بين تقنيات التعلم الآلي وتحليل الصور ومعالجة البيانات لمساعدة المزارعين في الحفاظ على صحة أشجارهم وتقليل خسائر المحصول. تقدم هذه الرسالة تصميم وتنفيذ وتقييم النظام، مما يدل على فعاليته في كشف الأمراض على فروع أشجار التفاح.

الكلمات الرئيسية: الذكاء الاصطناعي، كشف الأمراض، أشجار التفاح، تحليل الصور، التعلم الآلي، صحة المحاصيل، نيكتريا سينابرينا

RÉSUMÉ

Ce mémoire de master se concentre sur la détection des maladies sur les branches des pommiers en utilisant l'intelligence artificielle (IA). L'objectif de cette étude est de développer un système intelligent capable de détecter et de segmenter automatiquement la maladie de *Nectria cenabrina* sur les branches des pommiers à partir de l'analyse d'images. Les méthodes de détection manuelle traditionnelles sont longues, coûteuses et sujettes aux erreurs humaines. En exploitant les avancées de l'apprentissage automatique, du traitement d'image et des algorithmes d'apprentissage en profondeur, un système basé sur l'IA peut fournir des informations précises et en temps réel sur l'état de santé des pommiers, permettant aux agriculteurs de prendre des mesures préventives et d'optimiser leur production agricole. Le système proposé combine des techniques d'apprentissage automatique avec l'analyse d'images et le traitement de données pour aider les agriculteurs à maintenir la santé de leurs arbres et à réduire les pertes de récoltes. Ce mémoire présente la conception, la mise en œuvre et l'évaluation du système, démontrant son efficacité dans la détection des maladies sur les branches des pommiers.

Mots-clés : Intelligence Artificielle, Détection de Maladies, Pommiers, *Nectria cenabrina*, Analyse d'images, Apprentissage Automatique, Santé des Cultures

ABSTRACT

This master's thesis focuses on the detection of diseases on apple tree branches using artificial intelligence (AI). The aim of this study is to develop an intelligent system that can automatically detect and segment diseases on apple tree branches based on image analysis. Traditional manual detection methods are time-consuming, costly, and prone to human errors. By leveraging advancements in machine learning, image processing, and deep learning algorithms, an AI-based system can provide accurate and real-time information about the health status of apple trees, enabling farmers to take preventive measures and optimize their crop production. The proposed system combines machine learning techniques with image analysis and data processing to assist farmers in maintaining the health of their trees and reducing crop losses. This thesis presents the design, implementation, and evaluation of the system, demonstrating its effectiveness in disease detection on apple tree branches.

Keywords : Artificial Intelligence, Disease Detection, Apple Trees, Image Analysis, *Nectria cenabrina*, Machine Learning, Crop Health

TABLE DES MATIÈRES

Résum	i
Résumé	ii
Liste des figures	viii
Liste des tableaux	ix
Liste des abréviation	x
Introduction générale	1
Chapitre 1: Définition, impacts économiques et problèmes des pommiers	1
1.1 Introduction	1
1.2 Contexte et justification du travail	2
1.3 Valeur économique des pommiers au niveau mondiale	2
1.4 Valeur et impact économique en Algérie	3
1.4.1 Etude de cas : la province de Khenchela	5
1.5 Les maladies des pommiers	7
1.5.1 Feuilles et fruits	7
A Tavelure :	7

	B	Moniliose :	7
	C	Rouille grillagée :	8
1.5.2		La tige et les branches	8
	A	Chancre :	8
	B	Feu bactérien :	8
	C	Nectria cinnabarina :	9
	D	Maladie du collet :	9
	E	Pourriture brune de l'écorce :	9
	F	Maladie du corail :	10
	G	Chancre européen :	10
1.6		Les techniques de lutte contre les maladies des pommiers	10
1.7		Startup pommiers	11
1.8		Conclusion	12

Chapitre 2: Utilisation de l'intelligence artificielle (IA) dans le domaine de l'agriculture : Revue de littérature

			14
2.1		Introduction	14
2.2		IA et Pommiers	14
2.3		Deep Learning	15
	2.3.1	Réseaux neuronaux convolutifs (CNN)	16
	2.3.2	Réseaux de neurones récurrents (RNN)	17
	2.3.3	Réseaux de neurones générateurs adverses (GAN)	17
	2.3.4	Réseaux de neurones à mémoire à court terme (LSTM)	18
	2.3.5	Autoencodeurs	19
	2.3.6	Transformers	20
2.4		Machine learning	21
	2.4.1	Les machines à vecteurs de support (SVM)	21
	2.4.2	Les arbres de décision	22
	2.4.3	Gaussian Mixture Model (GMM)	24

2.5	Machine Learning vs Deep Learning	25
2.6	Travaux Connexes	26
2.7	Etude comparative	31
2.8	Conclusion	33
Chapitre 3: Conception et mise en œuvre du système		34
3.1	Introduction	34
3.2	Problématique	34
3.3	Architecture du système proposé	35
3.3.1	Acquisition	36
3.3.2	Prétraitement	36
A	Conversion de l'espace de couleur	37
B	Amélioration du contraste	38
C	Réduction du bruit	39
D	Seuillage et opérations morphologiques	40
E	Opération "et" bit à bit	41
3.3.3	Gaussian Mixture Model	42
A	Initialisation	44
B	Mise à jour des paramètres	45
3.3.4	Création du masque	47
3.3.5	Detection	48
3.4	Conclusion	48
Chapitre 4: Implémentation, Tests et Résultats		49
4.1	Introduction	49
4.2	Paramétrage	49
4.3	Outils Matériels	50
4.4	Outils Logiciels	51
4.4.1	Environnement de développement	51
4.4.2	Langage de programmation	51

4.4.3	Description des bibliothèques utilisées	51
4.5	La base de données	52
4.6	Les Métriques	53
4.7	Résultats Qualitatifs	55
4.8	Résultats Quantitatifs	60
4.9	Conclusion	62
	Conclusion générale	63
	Bibliographie	65

TABLE DES FIGURES

3.1	Schéma général du processus de détection de la maladie	35
3.2	Le processus de prétraitement proposé	37
3.3	pretraitement étape	42
3.4	Les étapes du GMM	43

LISTE DES TABLEAUX

1.1	Production mondiale de pommes par pays [Foo21]	4
1.2	Evolution de la culture du pommier dans la wilaya de kenchela (1998-2008) [Foo21]	5
2.1	Tableau comparatif des méthodes de détection des maladies dans les pommiers	32
4.1	Nombre d'images par variante d'arrière plan	53
4.2	Nombre d'images par variante de maladie	53
4.3	Quelques résultats de la detection de la maladie Nectria cenabrina dans les branche des pommiers	59
4.4	Résultats en F1-score, Recall et Précision dans quelques images de tests	61

LISTE DES ABRÉVIATION

IA : Intelligence artificielle

CNN : Réseaux neuronaux convolutifs

RNN : Réseaux de neurones récurrents

GAN : Réseaux de neurones générateurs adverses

LSTM : Réseaux de neurones à mémoire à court terme

ViT : Les vision transformers

SVM : Les machines à vecteurs de support

CART : Arbres de classification et de régression

GMM : Gaussian Mixture Model

ACP : Analyse des composantes principales

t-SNE : T-distributed stochastic neighbor embedding

FCM : L'algorithme de Fuzzy C-Means

RGA : L'algorithme de croissance de région

GA : L'algorithme génétique

CFS : La sélection de caractéristiques basée sur la corrélation

FPN : Réseaux pyramidaux de caractéristiques

ISResNet : Résiduels sélectifs améliorés

Faster RCNN : Réseaux de neurones convolutifs régionaux plus rapides

SKUAST-K : Université Sher-e-Kashmir des sciences et technologies agricoles Cachemire

RDF : Resource Description Framework

RGB : Red Green Blue

RVB : Rouge, vert et bleu

LAB : L = luminance , A et B : positionnement dans le spectre colorimétrique

CLAHE : Contrast Limited Adaptive Histogram Equalization

ML : Maximum de vraisemblance

EM : L'algorithme d'espérance-maximisation

INTRODUCTION GÉNÉRALE

Contexte et Problématique

L'intelligence artificielle (IA) a connu une croissance exponentielle ces dernières années dans de nombreux domaines, y compris l'agriculture. Dans le domaine spécifique de la détection des maladies des branches des pommiers, l'IA est essentielle pour améliorer la santé des arbres fruitiers, optimiser les rendements et garantir la qualité des récoltes.

Dans de nombreuses régions du monde, les pommiers sont une culture importante. Cependant, ces arbres peuvent être affectés par une variété de maladies, ce qui peut entraver leur croissance et leur productivité. La détection précoce et précise de ces maladies est essentielle pour prendre des mesures rapides et efficaces pour empêcher leur propagation et réduire les pertes pour les agriculteurs. Les experts agricoles ont traditionnellement détecté manuellement les maladies des branches des pommiers, un processus long, coûteux et sujet aux erreurs humaines. L'intelligence artificielle intervient dans cette situation. En utilisant les avancées de l'apprentissage automatique, du traitement d'image et des algorithmes d'apprentissage en profondeur, il est possible de développer des systèmes d'IA capables de détecter automatiquement et avec précision les signes de maladies sur les branches des pommiers.

Objectifs

Dans ce contexte, et face à cette problématique, L'objectif de ce projet de fin d'études (PFE) est de créer et de mettre en œuvre un système intelligent basé sur l'IA pour détecter les maladies des branches des pommiers. Ce système utilisera des techniques d'apprentissage automatique pour analyser des images de branches de pommiers afin d'identifier les anomalies ou les signes de maladies.

Ce système intelligent combinera les progrès de l'apprentissage automatique, du traitement d'image et de l'analyse des données pour fournir aux agriculteurs les outils nécessaires pour maintenir la santé de leurs arbres et optimiser leur production de pommes. Ce système permettra aux agriculteurs de prendre des mesures préventives et de traitement plus rapidement, réduisant les pertes de récolte et améliorant la durabilité de l'agriculture en fournissant des informations précises et en temps réel sur l'état de santé de leurs arbres.

Organisation du mémoire

Ce mémoire est organisé en quatre chapitres :

Le premier chapitre « **Définition, impacts économiques et problèmes des pommiers** » décrit l'importance des pommiers à l'échelle modelée, et leur impact économique et les maladies dangereuses qui affectent les pommiers.

Le deuxième chapitre « **Revue de littérature** » : Ce chapitre traite des méthodes de classification de DEEP et de machine learning, ainsi que des travaux liés à notre problématique.

Le troisième chapitre « **Conception et mise en œuvre du système** » : dans ce chapitre on discute La conception et l'implémentation d'un système de détection d'images de *Nectria cenabrina* , cette dernière comprend trois étapes : un prétraitement, entraînement de model , test.

dans le prétraitement on peut compter des etape principale comme : la Conversion d'espace de couleur, Seuillage et opérations morphologiques ,Opération "et" bit à bit .
pour l'entraînement de model : on prépare Les paramètres de model GMM .

Le quatrième chapitre « **Implémentation, Tests et Résultats**» explique comment le model GMM a été configuré, les outils utilisés, la découpe de la base de donnée, la mitrification et les résultats obtenus par le model.

CHAPITRE 1

DÉFINITION, IMPACTS ÉCONOMIQUES ET PROBLÈMES DES POMMIERS

1.1 Introduction

Les pommiers, appartenant à la famille des Rosaceae et au genre *Malus*, sont cultivés depuis des siècles pour leurs délicieux fruits, les pommes. Ils sont originaires d'Asie centrale, mais sont maintenant cultivés dans le monde entier dans une grande variété de climats. Les pommiers sont des arbres à feuilles caduques qui peuvent atteindre différentes tailles, allant de quelques mètres à plus de dix mètres de hauteur, en fonction de la variété. Ils ont une forme générale arrondie et se distinguent par leurs fleurs blanches ou roses au printemps, qui sont suivies par le développement des fruits. Les pommiers sont connus pour leur longévité et peuvent vivre plusieurs dizaines d'années. Ils nécessitent des conditions spécifiques pour se développer et produire des fruits de qualité, notamment une exposition adéquate au soleil, un sol bien drainé et une pollinisation croisée dans certains cas.

1.2 Contexte et justification du travail

Malheureusement, les pommiers sont également sujets à de nombreuses maladies, telles que la tavelure, la pourriture des fruits et diverses infections fongiques, qui peuvent réduire considérablement la productivité des vergers. La détection précoce de ces maladies revêt donc une importance cruciale pour la santé des pommiers et la durabilité de l'industrie fruitière. Dans cette optique, la recherche se concentre de plus en plus sur le développement de méthodes de détection automatique des maladies des pommiers, en utilisant des techniques telles que l'apprentissage automatique, l'imagerie et la vision par ordinateur, ainsi que l'analyse spectroscopique. Ces avancées prometteuses ouvrent de nouvelles perspectives pour la surveillance et la gestion efficace des maladies des pommiers, contribuant ainsi à améliorer la productivité et la durabilité de la culture de la pomme.

1.3 Valeur économique des pommiers au niveau mondiale

Les pommiers ont une valeur économique significative à l'échelle mondiale, contribuant à l'industrie agricole et alimentaire. Les pommes sont l'un des fruits les plus populaires et consommés dans de nombreux pays, ce qui en fait un produit de grande demande sur les marchés locaux et internationaux. La production mondiale de pommes dépasse 87.236.221 de tonnes chaque année, avec les principaux pays producteurs incluant la Chine, les États-Unis, la Russie, la Pologne et l'Allemagne. Les pommiers offrent également une diversité de variétés, allant des pommes sucrées et juteuses aux variétés plus acidulées et croquantes, répondant ainsi aux différents goûts et préférences des consommateurs. Les pommes sont consommées fraîches, transformées en jus, compotes, confitures, desserts et même utilisées dans la fabrication de produits dérivés tels que les cidres et les vinaigres de pomme. Cette diversité d'utilisation

confère aux pommiers une valeur économique considérable. De plus, l'industrie des pommiers génère des emplois à différents niveaux, de la production à la transformation, en passant par la distribution et la commercialisation. Les revenus générés par la culture des pommiers soutiennent les agriculteurs et les communautés rurales, favorisant ainsi le développement économique local. En outre, les pommiers contribuent à la préservation des paysages ruraux et à la promotion du tourisme dans les régions où la culture de la pomme est une activité majeure. En somme, la valeur économique des pommiers au niveau mondial est indéniable, et leur importance en tant que culture fruitière ne cesse de croître, tant pour les producteurs que pour les consommateurs.

Le pommier est le troisième fruit produit après les agrumes et la banane. Les deux variétés dominantes dans le monde sont les *Delicious rouges* et la *Golden Delicious* mais d'autres variétés ont connu une forte croissance, notamment la *Fuji* et la *Gala*.

Dans le monde, 87236221 tonnes de pommes sont produites par an selon [S1]. République populaire de Chine est le plus grand producteur de pommes au monde avec 42426578 tonnes de production par an. États-Unis d'Amérique arrive deuxième avec la production annuelle de 4997680 tonnes. République populaire de Chine et États-Unis d'Amérique produisent ensemble plus de 50 % du total mondial de pommes. Avec 3618752 tonnes de production par an, Turquie est le troisième producteur de pomme. Algérie, avec 558830 tonnes de production par an est classé à 23 [S1].

1.4 Valeur et impact économique en Algérie

Si la culture des pommes et des poires prédomine dans les pays au climat tempéré, l'Algérie a fait de gros efforts pour éliminer l'importation de ces deux fruits depuis l'indépendance à travers une bonne gestion des vergers, l'augmentation des rendements et l'augmentation des rendements. En 2007, il y avait uniquement 21 200 hectares de vergers de pommiers, principalement à Médéa, Batna, Tiaret, Blida et Khenchela. La même année, la production algérienne de pommes a atteint 181 000 tonnes avec un rendement de 85,3 Qx/Ha. Aujourd'hui, avec 558830 tonnes, l'Algérie est

N°	Pays	Production (tonnes)	Superficie (en hectares)	Rendement (Kg / Hectare)
01	Chine	42426578	2041197	20785,1
02	USA	4997680	118860	42046,8
03	Turquie	3618752	174439	20745,1
04	Pologne	3080600	166440	18508,8
05	Inde	2316000	308000	7519,5
06	Italie	2303690	55000	41885,3
07	Iran	2241124	100759	22242,4
08	Russie	1950800	210822	9253,3
09	France	1753500	50370	34812,4
10	Chili	1621321	32371	50085,6
11	Brésil	1222979	32405	37740,4
12	Ukraine	1153970	87700	13158,2
13	Ouzbékistan	1124017	98525	11408,4
14	Allemagne	991450	33980	29177,5
15	Afrique du Sud	891979	22373	39868,5
16	Maroc	809762	49731	16282,8
17	Corée du Nord	805965	70916	11365,1
18	Mexique	761483	52981	14372,8
19	Égypte	726794	28631	25384,9
20	Japon	701600	36000	19488,9
21	Espagne	638840	29640	21553,3
22	Moldavie	610921	51130	11948,4
23	Algérie	558830	32989	16939,9

TABLE 1.1 – Production mondiale de pommes par pays [Foo21]

classé en 23 position . Cette faible performance peut être attribuée à plusieurs causes dont :

- Le manque de connaissances liées à la biologie des espèces fruitières considérées.
- manque d'entretien du jardin.
- La non assimilation des techniques arboricoles modernes par les agriculteurs algériens (notamment les techniques d'élagage).
- Utilisation anarchique de porte-greffes et de variétés .
- La méconnaissance des zones propices à cette culture.
- Manque de connaissances sur l'application de la fertilisation. Dans notre pays,

cela se fait de manière empirique sans tenir compte des caractéristiques physico-chimiques du sol.

1.4.1 Etude de cas : la province de Khenchela

La province de Khenchela est l'une des principales régions productrices de pommes en Algérie, et la production concerne de nombreux districts (communes) de la province. Les vergers de pommiers sont principalement situés à Bouhmama, Kais, Chercher, El-Hamma et Babar.

A la fin de la campagne 1998/1999, la surface plantée en pommier était de 821 hectares avec un rendement de 38,377Qx . Depuis le démarrage du Programme National de Développement Agricole (P.N.D.A), la superficie plantée en pommiers dans la province de Khenchela a sensiblement augmenté passant de 2 598 hectares à 5 810 hectares , mais les rendements restent faibles et erratiques.

Année	Superficie (Ha)		Production (Qx)	Rendement (Qx/Ha)
	Totale	En rapport		
1998-1999	811	480	38 377	78
1999-2000	1022	521	35974	69,04
2000-2001	1518	563	56137	99,71
2001-2002	2087	730	56847	78
2002-2003	2598	793	90375	114
2003-2004	4304	990	94420	95,4
2004-2005	5352	1349	132280	98,1
2005-2006	5660	1584	225329	142,3
2006-2007	5605	1992	177877	89,3
2007-2008	5810	2452	220680	90

TABLE 1.2 – Evolution de la culture du pommier dans la wilaya de khenchela (1998-2008) [Foo21]

Lors de la saison agricole 2018/2019, la wilaya de Khenchela en Algérie s'est distinguée en devenant le leader de la filière pomiculture, avec une récolte de 1 466 000 quintaux de pommes, évaluée à une valeur de 25 milliards de dinars algériens.

Pour explorer de nouvelles méthodes, deux expériences pilotes de culture intensive de pommiers ont été lancées dans cette région, en utilisant la technique polonaise. Les vergers pilotes ont été implantés dans la commune de Bouhmama, où 12 000 arbres ont été plantés sur une superficie de quatre hectares, avec des variétés de pommes telles que Jeromine, Story et Granny Smith [S2].

Pour garantir le succès de ces expériences, diverses mesures ont été mises en place. Cela inclut l'importation de plants d'arbres en provenance d'Italie, l'installation de pollinisateurs, de filets anti-grêle, d'un système d'irrigation et l'adoption d'un programme de fertilisation polonais. On espère ainsi obtenir une production d'environ 200 quintaux par hectare dès la deuxième année, qui augmentera à 600 quintaux par hectare à partir de la troisième année[S2].

L'objectif principal de ces expériences est d'établir un partenariat fructueux entre les acteurs polonais et le conseil interprofessionnel de l'arboriculture fruitière, en vue d'augmenter la production, d'améliorer la qualité des fruits et de les rendre accessibles à un prix raisonnable. À terme, il est prévu d'adopter un programme d'exportation. Comparée à la méthode traditionnelle de pomiculture, la méthode polonaise permet de densifier les plantations en augmentant le nombre d'arbres par hectare (3 000 contre 600) et de commencer la récolte dès la deuxième année, évitant ainsi l'attente jusqu'à la cinquième année.

Cependant, il convient de souligner que malgré cette réussite, le rendement demeure relativement faible par rapport à la moyenne nationale, ce qui est attribué à une maîtrise insuffisante des facteurs de production et à une demande limitée pour cette culture. En plus, l'expansion de cette activité opportuniste a conduit à l'émergence de certains ravageurs spécifiques qui réduisent le rendement et la qualité des fruits. En plus de ces problèmes, il existe plusieurs autres facteurs qui affectent négativement la production, nous mentionnons en particulier, l'installation de cultures intercalaires, une pratique ancienne d'élagage, utilisant des systèmes d'irrigation traditionnels qui ne répondent pas à la demande en eau. Niveaux de plantation très élevés dans la zone, dont la croissance a été importante au cours des dernières années.

1.5 Les maladies des pommiers

Les pommiers sont sensibles à plusieurs maladies qui peuvent affecter leur croissance et leur production de fruits.

1.5.1 Feuilles et fruits

A Tavelure :

La tavelure du pommier et du poirier, causée par le champignon *Venturia inaequalis*, est l'une des maladies les plus répandues et destructrices pour ces arbres fruitiers. Les spores de ce champignon se propagent principalement par le vent et l'eau, infectant les feuilles, les fleurs et les fruits. Les premiers symptômes apparaissent sous la forme de taches vert olive sur la face supérieure des feuilles, qui se transforment ensuite en taches brunes avec un aspect velouté. Si la maladie n'est pas contrôlée, elle peut causer la chute prématurée des feuilles et réduire considérablement la qualité et la quantité de la récolte [Jam+09].

B Moniliose :

La moniliose, également connue sous le nom de pourriture des fruits ou pourriture brune, est causée par les champignons du genre *Monilinia*, notamment *Monilinia fructigena* et *Monilinia laxa*. Ces champignons hibernent dans les fruits momifiés restés sur l'arbre ou au sol. Les spores se dispersent par le vent, la pluie et les insectes, infectant les fleurs, les jeunes fruits et même les fruits récoltés. Les symptômes de la moniliose comprennent des taches brunes entourées d'une zone de décoloration jaune sur les fruits, suivies d'un dessèchement du fruit sans qu'il ne tombe [Per12].

C Rouille grillagée :

La rouille grillagée est une maladie fongique causée par le champignon *Gymnosporangium sabinae*. Ce champignon a un cycle de vie complexe, nécessitant deux hôtes pour compléter son cycle : les pommiers et les genévriers. Les spores de la rouille grillagée se propagent par le vent et infectent les feuilles de pommiers, provoquant l'apparition de taches jaune orangé qui évoluent en excroissances linéaires. Cette maladie peut entraîner la défoliation précoce des arbres et réduire la production de fruits [SA22].

1.5.2 La tige et les branches

Les maladies des tiges et des branches des pommiers peuvent avoir des conséquences dévastatrices sur la santé des arbres et leur production de fruits. Voici quelques-unes des maladies les plus courantes qui affectent les tiges et les branches des pommiers :

A Chancre :

Le chancre est une maladie qui présente un risque accru pendant les années de fortes précipitations. Elle se manifeste par des lésions sur l'écorce du tronc ou des branches, exposant ainsi le bois à l'air. Des pustules contenant des spores de couleur rouge ou blanche peuvent parfois se former à la surface des blessures. Le chancre affaiblit l'arbre et peut entraîner la pourriture du bois, compromettant ainsi sa stabilité et sa capacité à produire des fruits [Per77].

B Feu bactérien :

Bien que le feu bactérien soit plus fréquent chez les poiriers, les pommiers peuvent également être touchés par cette maladie grave. Les symptômes typiques du feu bactérien comprennent la flétrissure des fleurs, la nécrose des branches, un écoulement blanc à doré sur l'écorce ou les fruits, ainsi que la formation d'ulcères sur les branches

infectées. Si la maladie n'est pas traitée rapidement, elle peut entraîner la mort de l'arbre. Les bactéries responsables du feu bactérien se développent préférentiellement dans des conditions chaudes et humides. Il est crucial de signaler la présence de cette maladie aux autorités municipales, car elle est réglementée et nécessite une gestion appropriée [Lec+01].

C Nectria cinnabarina :

Cette maladie est facilement identifiable grâce à l'apparition de pustules de couleur orange, rose ou rouge brique sur l'écorce des arbres. Les spores responsables de la maladie pénètrent ensuite dans les plaies des arbres affaiblis, provoquant la formation de chancres. La maladie affaiblit l'arbre et peut entraîner la pourriture du bois, compromettant ainsi sa santé globale et sa capacité à produire des fruits [DLS95].

D Maladie du collet :

Cette maladie affecte la zone de transition entre la tige principale et les racines, appelée collet. Elle est généralement causée par des champignons du genre *Phytophthora* ou des bactéries du genre *Agrobacterium*. Les symptômes comprennent un flétrissement soudain, un jaunissement des feuilles et un dépérissement de l'arbre. Les champignons ou les bactéries attaquent les tissus du collet, entraînant une pourriture qui affecte la circulation des nutriments et de l'eau dans l'arbre [USS01].

E Pourriture brune de l'écorce :

Cette maladie, causée par le champignon *Phomopsis*, provoque la pourriture de l'écorce et des tissus sous-jacents de la tige et des branches. Les symptômes incluent l'apparition de taches brunes ou noires sur l'écorce, qui s'étendent progressivement en formant des zones de pourriture. Les branches touchées peuvent se flétrir et mourir, entraînant une réduction de la vigueur de l'arbre et de sa capacité à produire des fruits [Dan92].

F Maladie du corail :

Cette maladie, également connue sous le nom de maladie du corail du pommier, est causée par le champignon *Nectria galligena*. Elle se manifeste par la formation de chancres corail, des excroissances en forme de corail de couleur rose à orange vif, sur l'écorce des tiges et des branches. Ces chancres entraînent la mort des tissus, affaiblissant l'arbre et pouvant entraîner des déformations, des nécroses et la réduction de la production de fruits [HM14].

G Chancre européen :

Cette maladie est causée par le champignon *Neonectria ditissima*. Elle se caractérise par la formation de chancres sur les tiges et les branches des pommiers. Les chancres ont une apparence brun foncé avec des bords en relief et peuvent provoquer la mort des tissus infectés. La maladie peut entraîner un affaiblissement de l'arbre, une diminution de la production de fruits et une réduction de la longévité de l'arbre [DG+20].

1.6 Les techniques de lutte contre les maladies des pommiers

Il existe plusieurs techniques de lutte contre les maladies des pommiers. Voici quelques-unes des méthodes couramment utilisées :

- Pratiques culturales appropriées : Maintenir une bonne santé globale des arbres en adoptant des pratiques culturales appropriées est essentiel pour prévenir les maladies. Cela inclut la taille régulière des arbres pour favoriser une bonne circulation de l'air, l'élimination et la destruction des débris végétaux infectés, la rotation des cultures et la gestion adéquate de l'irrigation.
- Utilisation de variétés résistantes : Le choix de variétés de pommiers résistantes aux maladies spécifiques peut contribuer à réduire l'incidence des infections.

Certaines variétés sont génétiquement plus résistantes à certaines maladies, ce qui peut réduire le besoin de traitements chimiques.

- Traitement chimique : Dans les cas où les mesures préventives ne suffisent pas, l'utilisation de produits phytosanitaires peut être nécessaire pour contrôler les maladies. Il est important de choisir des produits adaptés aux maladies ciblées, de suivre attentivement les instructions d'utilisation et de respecter les délais de sécurité avant la récolte.
- Élagage et élimination des parties infectées : L'élimination des parties infectées des arbres, telles que les branches atteintes de chancre, peut aider à réduire la propagation des maladies. Il est important de désinfecter les outils de coupe entre chaque arbre pour éviter la transmission de maladies d'un arbre à l'autre.
- Surveillance régulière : Une surveillance régulière des vergers permet de détecter rapidement les signes de maladies et d'intervenir rapidement. Il est recommandé d'inspecter les arbres régulièrement pour repérer les symptômes tels que des lésions, des taches ou des déformations sur les tiges, les branches ou les feuilles.

1.7 Startup pommiers

Les startups axées sur les pommiers sont des entreprises innovantes qui se spécialisent dans divers aspects de la culture des pommiers et de la production de pommes. Voici quelques exemples de startups dans ce domaine :

- Treedom : Treedom est une startup qui propose un modèle de plantation d'arbres en ligne. Ils permettent aux particuliers et aux entreprises de planter des pommiers (et d'autres types d'arbres fruitiers) à distance et de suivre leur croissance grâce à une plateforme en ligne. Cela favorise la durabilité environnementale et permet aux utilisateurs de soutenir des projets agricoles locaux.
- Phytoponics : Phytoponics est une startup qui développe des systèmes hydroponiques innovants pour la culture des pommiers et d'autres cultures. Leur

technologie permet de cultiver des arbres fruitiers en utilisant moins d'eau et d'espace, tout en augmentant les rendements. Ils proposent des solutions pour les vergers urbains et les serres commerciales.

- DroneAG : DroneAG est une startup qui utilise des drones équipés de capteurs pour surveiller les vergers de pommiers et détecter les maladies précoces ou les problèmes de croissance. Leurs drones collectent des données et fournissent des analyses en temps réel, permettant aux agriculteurs de prendre des mesures préventives et d'optimiser leurs pratiques de culture.
- POMIO : POMIO est une startup qui propose une plateforme en ligne pour la gestion des vergers de pommiers. Leur logiciel permet aux agriculteurs de suivre et de gérer les tâches liées à la culture des pommiers, telles que la fertilisation, l'irrigation, la protection contre les maladies et la planification des récoltes. Ils offrent également des conseils et des recommandations basés sur l'analyse des données.

Ces startups cherchent à apporter des innovations et des solutions technologiques à l'industrie de la culture des pommiers, en améliorant la productivité, la durabilité et la rentabilité pour les agriculteurs. Elles contribuent également à l'avancement des pratiques agricoles respectueuses de l'environnement et à la promotion d'une alimentation saine et durable.

1.8 Conclusion

Dans ce chapitre, nous avons examiné les pommiers et leur impact économique en Algérie et dans le monde entier. Toutefois, ces arbres sont également vulnérables à diverses maladies qui peuvent affecter leurs feuilles, leurs tiges et leurs branches.

Il est crucial de mettre en œuvre des mesures préventives et de contrôle pour atténuer l'impact de ces maladies. Cela implique l'adoption de bonnes pratiques culturales telles que la taille et l'élimination des tissus infectés, l'utilisation de fongicides ou de bactéricides recommandés, ainsi que la création d'un environnement sain pour

les pommiers. Une surveillance régulière des arbres et le recours à des experts en arboriculture peuvent également contribuer à prévenir la propagation et la gravité de ces maladies.

Les startups offrent des solutions prometteuses en apportant des innovations et des solutions technologiques à l'industrie de la culture des pommiers. Elles visent à améliorer la productivité, la durabilité et la rentabilité pour les agriculteurs. De plus, ces startups favorisent l'adoption de pratiques agricoles respectueuses de l'environnement et soutiennent une alimentation saine et durable.

CHAPITRE 2

UTILISATION DE L'INTELLIGENCE ARTIFICIELLE (IA) DANS LE DOMAINE DE L'AGRICULTURE : REVUE DE LITTÉRATURE

2.1 Introduction

L'utilisation de l'intelligence artificielle (IA) dans le domaine de l'agriculture a ouvert de nouvelles perspectives et opportunités passionnantes. Grâce à l'IA, les agriculteurs peuvent désormais bénéficier de solutions technologiques avancées pour améliorer l'efficacité, la productivité et la durabilité de leurs exploitations. Cette convergence entre l'agriculture et l'IA a le potentiel de transformer l'industrie agricole et de répondre aux défis croissants liés à la sécurité alimentaire, aux ressources limitées et aux changements climatiques.

2.2 IA et Pommiers

Dans le contexte spécifique des pommiers, l'utilisation de l'IA a également suscité un intérêt considérable. Les pommiers sont des cultures clés dans de nombreux pays,

et leur santé et leur productivité sont essentielles pour répondre à la demande croissante de pommes de qualité. L'IA offre des opportunités pour surveiller, analyser et optimiser les performances des pommiers de manière précise et efficace.

Grâce à l'IA, les agriculteurs peuvent utiliser des systèmes de surveillance et de détection avancés pour identifier précocement les maladies, les ravageurs ou les carences nutritionnelles dans les pommiers. Les algorithmes d'IA peuvent analyser les données collectées à partir de capteurs, d'imagerie par satellite et d'autres sources pour fournir des informations en temps réel sur l'état de santé des pommiers. Cela permet aux agriculteurs de prendre des décisions éclairées et de mettre en place des mesures de prévention ou de traitement plus ciblées, réduisant ainsi les pertes de récolte et les coûts liés aux intrants agricoles.

En outre, l'IA peut également contribuer à l'optimisation de la gestion des ressources, telles que l'irrigation et la fertilisation, en fonction des besoins spécifiques des pommiers. Les modèles d'IA peuvent analyser les données historiques, les conditions météorologiques actuelles et les prévisions pour recommander des stratégies de gestion adaptées. Cela permet d'améliorer l'utilisation des ressources, de réduire les impacts environnementaux et d'optimiser la croissance et la qualité des pommiers.

2.3 Deep Learning

La littérature présente plusieurs méthodes basées sur le Deep Learning, qui fournissent des capacités avancées pour résoudre des problèmes complexes dans divers domaines tels que la vision par ordinateur, le traitement du langage naturel, la reconnaissance de la parole et bien d'autres. Chaque méthode possède ses propres caractéristiques et avantages, ce qui les rend adaptées à différentes tâches et types de données. Dans cette section, nous allons mettre en évidence les méthodes les plus puissantes.

2.3.1 Réseaux neuronaux convolutifs (CNN)

Un réseau neuronal convolutif (CNN) est un type de réseau neuronal profond qui incorpore différentes techniques d'apprentissage en profondeur. Les CNN sont des variantes des réseaux de neurones multicouches, avec des couches supplémentaires spécifiquement conçues pour réduire la dimensionnalité des données et le nombre de paramètres. Ils sont couramment utilisés pour la détection d'objets, la classification d'images et la régression, permettant ainsi l'extraction de caractéristiques. Les CNN présentent plusieurs avantages, notamment leur capacité à effectuer des convolutions sur de grands ensembles de données à grande vitesse pendant la phase d'apprentissage. Dans les prochaines sections, nous fournirons un aperçu des réseaux neuronaux convolutifs, en décrivant leurs composants essentiels tels que les couches convolutives, les couches de pooling et les couches entièrement connectées. Cette architecture unique rend les CNN particulièrement adaptés aux tâches de classification d'images, permettant ainsi un apprentissage rapide des réseaux multicouches, qui se distinguent dans la classification d'images[CR93].

En ce qui concerne l'architecture, un réseau neuronal convolutif (CNN) est composé d'une couche d'entrée, de couches cachées et d'une couche de sortie. Les couches intermédiaires du CNN sont communément appelées "cachées", car elles cachent leurs entrées et leurs sorties grâce à l'utilisation de fonctions d'activation et de convolutions. Dans un CNN, les couches cachées sont constituées de couches qui effectuent des convolutions, suivies d'autres couches telles que des couches de pooling et des couches entièrement connectées.

Les composants clés d'un CNN comprennent les noyaux (kernels) et les filtres, la couche convolutive, la couche de pooling, le processus d'aplanissement (flattening) et la couche entièrement connectée. Les noyaux et les filtres sont utilisés pour extraire des caractéristiques spécifiques de l'image en effectuant des opérations de convolution. La couche convolutive applique ces noyaux et filtres pour détecter des motifs et des caractéristiques au sein de l'image. Ensuite, la couche de pooling réduit la taille

spatiale des données en extrayant les caractéristiques les plus importantes. Le processus d'aplanissement convertit les données en un vecteur unidimensionnel, qui est ensuite alimenté dans les couches entièrement connectées pour effectuer la classification finale [CR93].

2.3.2 Réseaux de neurones récurrents (RNN)

Les RNN, ou Réseaux de Neurones Récurrents, sont des modèles largement utilisés pour traiter des données séquentielles, qu'il s'agisse de séquences de mots dans le traitement du langage naturel ou de séries temporelles dans l'analyse des données chronologiques. La principale caractéristique des RNN est leur capacité à conserver une mémoire interne qui leur permet de capturer et de mémoriser les informations des états précédents. Cette mémoire interne est essentielle pour comprendre le contexte et les dépendances temporelles des données séquentielles. Grâce à cette capacité, les RNN peuvent apprendre à prédire, générer ou classer des données séquentielles en prenant en compte les relations et les motifs qui se produisent au fil du temps. Cela en fait des outils puissants pour des applications telles que la traduction automatique, la génération de texte, la reconnaissance de la parole et bien d'autres où la compréhension du contexte est primordiale [Yin+17].

2.3.3 Réseaux de neurones générateurs adverses (GAN)

Les GAN, ou Réseaux Génératifs Antagonistes, sont des modèles puissants utilisés pour générer des données réalistes à partir d'un ensemble d'apprentissage donné. Ils sont composés de deux réseaux neuronaux en compétition : un générateur et un discriminateur. Le générateur a pour objectif de créer de nouvelles données qui ressemblent à celles de l'ensemble d'apprentissage, tandis que le discriminateur cherche à distinguer les données réelles de celles générées par le générateur.

Cette compétition entre le générateur et le discriminateur crée un processus d'apprentissage itératif où les deux réseaux s'améliorent mutuellement. Le générateur apprend à générer des données de plus en plus réalistes en fonction des informations fournies par le discriminateur, tandis que le discriminateur apprend à devenir plus précis dans la distinction entre les données réelles et générées.

Les GAN sont couramment utilisés dans la génération d'images réalistes, où ils peuvent créer des visages, des paysages et d'autres objets visuels convaincants. Ils sont également utilisés dans d'autres domaines de création de contenu, tels que la synthèse de voix, où ils peuvent générer des voix humaines naturelles et expressives à partir de données d'apprentissage.

Grâce à leur capacité à générer des données réalistes et de haute qualité, les GAN ont ouvert de nouvelles possibilités dans la création artistique, la réalité virtuelle, la simulation de données et d'autres domaines où la génération de contenu réaliste est essentielle [Pan+19].

2.3.4 Réseaux de neurones à mémoire à court terme (LSTM)

Les LSTM, ou Réseaux de Neurones à Mémoire à Long Terme, sont une variante des RNN qui ont été spécialement conçus pour résoudre le problème de l'apprentissage à long terme des dépendances dans les séquences. Contrairement aux RNN traditionnels, qui peuvent avoir du mal à capturer les dépendances à long terme en raison de l'effet de disparition du gradient, les LSTM sont dotés d'une architecture spéciale qui leur permet de mémoriser et d'utiliser des informations sur de longues périodes de temps.

Cette capacité des LSTM à conserver une mémoire à long terme les rend extrêmement précieux dans des domaines tels que le traitement du langage naturel, la traduction automatique et d'autres tâches impliquant des séquences complexes. Par exemple, lors de la modélisation du langage naturel, les LSTM peuvent apprendre les structures grammaticales et les relations entre les mots sur de longues distances, ce

qui leur permet de générer du texte cohérent et de prédire la probabilité de séquences de mots.

Les LSTM utilisent des unités de mémoire spéciales appelées "cellules" qui permettent de stocker et d'accéder à des informations sur de longues périodes de temps. Ces cellules sont dotées de portes qui contrôlent le flux d'informations dans et hors de la cellule, permettant ainsi aux LSTM de décider quand mémoriser ou oublier des informations.

En raison de leur capacité à capturer les dépendances à long terme, les LSTM sont largement utilisés dans des applications telles que la génération de texte, la prédiction de séquences, la reconnaissance de la parole et d'autres tâches où la compréhension des contextes complexes est essentielle. Leur architecture avancée et leur capacité à gérer efficacement les séquences font des LSTM des outils précieux pour résoudre des problèmes de traitement des séquences dans des domaines variés [SJ19].

2.3.5 Autoencodeurs

Les autoencodeurs sont des architectures de réseaux neuronaux qui jouent un rôle essentiel dans l'apprentissage de représentations comprimées de données non étiquetées. Leur objectif principal est de capturer les caractéristiques les plus importantes des données en réduisant leur dimensionnalité, tout en préservant autant d'informations que possible.

Les autoencodeurs se composent de deux parties principales : l'encodeur et le décodeur. L'encodeur prend les données d'entrée et les transforme en une représentation latente de dimension réduite. Cette représentation latente contient des informations essentielles sur les données, tout en étant de taille réduite par rapport à l'entrée d'origine. Le décodeur, quant à lui, prend la représentation latente et tente de reconstruire les données d'origine à partir de cette représentation[BKG20].

L'idée fondamentale derrière les autoencodeurs est que le processus d'encodage et de décodage force le réseau à apprendre une représentation comprimée qui capture les structures et les caractéristiques les plus importantes des données. En réduisant

la dimensionnalité, les autoencodeurs peuvent extraire les informations essentielles tout en éliminant le bruit et les détails superflus.

Les autoencodeurs sont largement utilisés dans diverses applications. Dans le domaine de la compression d'images, par exemple, les autoencodeurs peuvent réduire la taille des images en apprenant une représentation latente compacte. Dans la détection d'anomalies, les autoencodeurs peuvent apprendre à reconstruire des données normales et identifier les points de données qui s'écartent de la reconstruction attendue, ce qui permet de détecter des comportements anormaux[Pin+20].

En outre, les autoencodeurs sont également utilisés dans le prétraitement des données, notamment dans la réduction de dimensionnalité, la suppression du bruit, l'augmentation de données et d'autres techniques visant à améliorer la qualité des données d'entrée avant de les utiliser dans d'autres tâches d'apprentissage automatique.

Grâce à leur capacité à apprendre des représentations comprimées et informatives des données, les autoencodeurs sont des outils puissants pour la compréhension et la transformation des données non étiquetées. Leur utilisation permet d'améliorer l'efficacité et la performance des systèmes d'apprentissage automatique dans diverses applications [Pin+20; BKG20].

2.3.6 Transformers

Les transformers sont une architecture révolutionnaire dans le domaine du deep learning, qui a connu un succès retentissant dans de nombreuses tâches liées au traitement du langage naturel. Contrairement aux modèles basés sur les réseaux récurrents ou convolutifs, les transformers utilisent une approche entièrement basée sur l'attention pour capturer les dépendances à longue distance dans les séquences [Lin+22].

L'élément clé des transformers est le mécanisme d'attention, qui permet au modèle de pondérer l'importance de chaque élément de la séquence en fonction de sa relation avec les autres éléments. Cela permet aux transformers de prendre en compte simultanément toutes les parties de la séquence lors de la prise de décision, ce qui améliore considérablement leur capacité à comprendre le contexte global [Lin+22].

Les transformers ont été largement adoptés dans des tâches telles que la traduction automatique, la génération de texte, la résumé automatique, la reconnaissance d'entités nommées et bien d'autres. Leur architecture modulaire et parallélisable les rend particulièrement efficaces pour le traitement de grandes quantités de données et l'entraînement sur des ressources de calcul distribuées [Kha+22].

Un autre avantage des transformers est leur capacité à apprendre des représentations sémantiques riches. Grâce à leur architecture attentionnelle, ils peuvent capturer des relations complexes entre les mots et les phrases, ce qui facilite l'apprentissage de modèles de langage de haute qualité et leur transfert vers d'autres tâches.

Les transformers ont également été étendus à d'autres domaines, tels que la vision par ordinateur, avec des modèles tels que les vision transformers (ViT). Ces modèles ont montré des performances impressionnantes dans des tâches de classification d'images et de segmentation, en exploitant également les avantages de l'attention pour traiter les informations visuelles. Les transformers ont également été étendus à d'autres domaines, tels que la vision par ordinateur, avec des modèles tels que les vision transformers (ViT). Ces modèles ont montré des performances impressionnantes dans des tâches de classification d'images et de segmentation, en exploitant également les avantages de l'attention pour traiter les informations visuelles [Kha+22].

2.4 Machine learning

2.4.1 Les machines à vecteurs de support (SVM)

Les Machines à Vecteurs de Support (SVM) sont des algorithmes puissants utilisés dans divers domaines tels que la reconnaissance de formes, la classification et l'apprentissage automatique. Elles se basent sur la représentation des données sous forme de vecteurs réels dans un espace multidimensionnel. L'objectif des SVM est de trouver une frontière de décision optimale qui sépare les données en classes distinctes [VM02].

Contrairement à certaines méthodes de classification, les SVM ne cherchent pas à estimer directement la densité de probabilité des données, mais se concentrent sur la recherche d'une fonction de décision qui maximise la marge entre les différentes classes. Cette fonction de décision doit attribuer correctement les étiquettes de classe en satisfaisant l'équation

$$\text{sign}(f(x)) = y_i \quad (2.1)$$

où y_i représente la classe d'appartenance attendue.

Ce qui distingue les SVM des autres approches est leur critère d'optimalité, qui vise à trouver la frontière de décision la plus robuste et la plus généralisable possible. Les SVM s'appuient sur des techniques avancées de classification supervisée et ont été développées en tirant parti des travaux antérieurs dans le domaine.

Les SVM offrent plusieurs avantages, tels que leur capacité à gérer des ensembles de données complexes et à travailler efficacement dans des espaces de grande dimension. Elles peuvent également gérer des données non linéairement séparables en utilisant des techniques de noyau pour projeter les données dans des espaces de dimension supérieure. Cela permet aux SVM de capturer des relations non linéaires entre les variables et d'améliorer la précision de la classification [CB21].

2.4.2 Les arbres de décision

La construction d'arbres de décision à partir de données est une discipline ancienne qui a été initiée par Morgan et Sonquist en 1963 avec l'utilisation des arbres de régression dans un processus de prédiction et d'explication. Cette approche a donné naissance à une famille de méthodes reposant sur la représentation par arbres, qui ont été étendues pour aborder des problèmes de discrimination et de classement. Parmi ces méthodes, on retrouve les approches THAID et CHAID. L'apogée de cette approche est représentée par la méthode CART de Breiman et al., qui est largement documentée dans une monographie encore considérée comme une référence aujourd'hui [De 13].

Dans le domaine de l'apprentissage automatique, de nombreux travaux se sont appuyés sur la théorie de l'information. L'algorithme ID3 de Quinlan a été particulièrement influent dans les années 80, avec de nombreuses publications. Son travail a connu un tournant important dans les années 90 avec l'introduction de la méthode C4.5. Certains travaux, tels que ceux de Bourroche et Tenenhaus avec la méthode ELISEE, ont adopté une approche statistique, tandis que les travaux de Picard sur les pseudo-questionnaires se rapprochent davantage de la théorie de l'information. Ces différentes approches ont contribué à l'émergence du concept de graphes latticiels, qui a été popularisé par les graphes d'induction avec la méthode SIPINA.

En ce qui concerne les algorithmes d'arbre de décision, il existe plusieurs variantes notables :

- ID3 : est considéré comme l'un des premiers et utilise le gain d'informations comme critère de division. Il s'arrête de se développer lorsque toutes les instances appartiennent à une seule valeur de sortie ou lorsque le gain d'informations maximal est nul. Cependant, ID3 ne prend pas en charge les attributs numériques ou les valeurs manquantes, et il n'applique pas de procédure d'élagage [JDF09].
- C4.5 : c'est une amélioration de l'algorithme ID3. Il utilise le rapport de gain comme critère de division et s'arrête lorsque le nombre d'instances à diviser est inférieur à un seuil prédéfini. Après la phase de croissance, l'élagage basé sur les erreurs est effectué. Contrairement à ID3, C4.5 est capable de gérer les attributs numériques et les valeurs manquantes en utilisant des critères de rapport de gain corrigés [Hss+14; Hss+14].
- CART : est caractérisé par la construction d'arbres binaires, où chaque nœud interne possède exactement deux branches sortantes. Les divisions sont sélectionnées à l'aide des critères de Gini ou de "Twoing". Une particularité importante de CART est sa capacité à générer des arbres de régression, où les

feuilles prédisent une valeur numérique plutôt qu'une classe. Pour la régression, CART recherche les divisions qui minimisent l'erreur quadratique de prédiction. La prédiction dans chaque feuille est basée sur la moyenne pondérée des valeurs du nœud [Azi+19].

2.4.3 Gaussian Mixture Model (GMM)

Les modèles de mélange gaussien (GMM) sont largement utilisés dans divers domaines tels que la vision par ordinateur, la reconnaissance vocale et la finance, offrant ainsi une gamme étendue d'applications. Ils sont utilisés pour des tâches telles que la segmentation d'images, la détection d'anomalies et la modélisation de données financières. Par exemple, les GMM peuvent être appliqués pour identifier différents types d'objets dans une image ou détecter des transactions frauduleuses dans des données financières [BBW01].

L'approche de clustering doux utilisée par les modèles de mélange gaussien diffère du clustering dur en attribuant à chaque point de données une probabilité d'appartenance à chaque cluster, plutôt que de l'assigner à un seul cluster. Cela permet aux points de données d'appartenir simultanément à plusieurs groupes, avec des degrés d'appartenance variables. En revanche, le clustering dur attribue chaque point de données à un seul groupe en se basant sur une mesure de distance ou de similarité [BBW01].

Bien que les modèles de mélange gaussien soient principalement utilisés pour le regroupement, ils peuvent également être employés pour la réduction de la dimensionnalité. Une approche consiste à modéliser la distribution des données à l'aide d'un GMM, puis à générer de nouveaux points de données de dimension inférieure en échantillonnant le modèle et en projetant les résultats dans un espace de dimension réduite. Cependant, il convient de noter que d'autres techniques, telles que l'analyse en composantes principales (ACP) et le t-SNE, sont plus couramment utilisées pour la réduction de la dimensionnalité [ZA14].

2.5 Machine Learning vs Deep Learning

Le machine learning (apprentissage automatique) et le deep learning (apprentissage profond) sont deux sous-domaines de l'intelligence artificielle qui ont des approches distinctes mais complémentaires.

Le machine learning se concentre sur l'utilisation d'algorithmes pour analyser et interpréter les données, afin de prendre des décisions ou de faire des prédictions. Il repose sur la création de modèles statistiques qui sont entraînés sur des ensembles de données pour reconnaître des schémas et des relations. Les avantages du machine learning résident dans sa polyvalence et sa capacité à traiter différents types de données, de la classification à la régression en passant par le regroupement. Il peut être utilisé efficacement avec des ensembles de données de taille relativement petite et ne nécessite pas toujours un grand nombre de données pour fournir des résultats précis. Cependant, il peut être limité par sa capacité à extraire des caractéristiques complexes et à capturer des relations non linéaires entre les variables [SS18].

Le deep learning, quant à lui, est une branche du machine learning qui s'appuie sur des réseaux de neurones artificiels profonds. Ces réseaux sont composés de nombreuses couches de neurones interconnectés, ce qui leur permet d'apprendre des représentations hiérarchiques des données. Le deep learning excelle dans la reconnaissance de modèles complexes et dans la capture de relations non linéaires, ce qui en fait une approche puissante pour la vision par ordinateur, le traitement du langage naturel et d'autres tâches nécessitant une compréhension approfondie des données. Cependant, le deep learning nécessite généralement de grandes quantités de données d'entraînement et de puissance de calcul pour former efficacement des réseaux de neurones profonds. Il peut également être plus difficile à interpréter en raison de sa complexité [SS18].

2.6 Travaux Connexes

Au cours des dernières années, la recherche s'est penchée sur la détection des maladies dans les cultures et les plantations, et cette attention s'est également étendue à la détection précoce des maladies chez les arbres. Cette détection précoce revêt une importance cruciale pour préserver la santé des écosystèmes. Elle repose sur une combinaison de méthodes visuelles, d'observations sur le terrain et de techniques d'analyse avancées.

Les arboriculteurs et les forestiers jouent un rôle essentiel dans la détection précoce des maladies chez les arbres, car ils sont souvent les premiers à repérer les signes visibles de ces affections. Leurs observations attentives peuvent révéler des symptômes tels que des taches, des dépérissements, des nécroses ou des déformations. Ils peuvent également identifier des indicateurs tels que des écoulements de résine, des chancres ou des écorces fissurées, qui les mettent sur la piste de potentielles maladies.

En plus des méthodes traditionnelles basées sur les observations visuelles, les techniques de diagnostic avancées jouent un rôle de plus en plus important. L'intelligence artificielle (IA) est notamment utilisée pour la détection des maladies chez les arbres. Grâce à l'IA, il est possible de développer des systèmes automatisés capables d'analyser des images et de détecter des signes de maladies de manière rapide et précise.

Dans la littérature scientifique, plusieurs articles ont abordé cette problématique en proposant des méthodes innovantes pour la détection automatique des maladies chez les arbres. Ces travaux exploitent l'utilisation de techniques d'apprentissage automatique et de réseaux de neurones pour entraîner des modèles capables de reconnaître les symptômes caractéristiques des maladies.

Article 1 : A Novel Convolutional Neural Network Based Model for Recognition and Classification of Apple Leaf Diseases L'article présente un modèle qui utilise un réseau de neurones convolutifs (CNN) pour la reconnaissance et la classification

des maladies des feuilles de pommier. Le modèle se compose de quatre étapes principales : le prétraitement des données, la segmentation, l'extraction des caractéristiques et la classification.

La première étape consiste à améliorer le contraste des images en utilisant une technique d'étirement du contraste afin de mettre en évidence les informations masquées. Ensuite, la segmentation est réalisée en utilisant l'algorithme de clustering Fuzzy C-Means (FCM) pour identifier la zone affectée sur la feuille.

Dans la troisième étape, le modèle CNN est utilisé pour extraire les caractéristiques distinctives qui permettent d'identifier les modèles et les particularités des feuilles malades. Enfin, le modèle CNN est utilisé pour classer l'image d'entrée comme étant saine ou malade, en se basant sur les informations extraites précédemment.

Le modèle a été entraîné et validé sur un ensemble de données composé de 400 échantillons d'images, dont 200 feuilles de pommier saines et 200 feuilles malades. Bien que le modèle ait été spécifiquement conçu pour les maladies des feuilles de pommier, les techniques utilisées, telles que le prétraitement par étirement du contraste et la segmentation par clustering Fuzzy C-Means, peuvent également être appliquées à d'autres types de plantes.

De plus, le modèle CNN peut être entraîné sur des ensembles de données d'autres maladies des plantes, permettant ainsi de les reconnaître et de les classer avec précision. L'objectif principal de cette approche est donc de fournir un outil automatisé pour la détection précoce et la classification des maladies foliaires du pommier, tout en ouvrant la voie à des applications similaires dans d'autres contextes [YY+20].

Article 2 : Disease Detection in Apple Leaves Using Image Processing Techniques.

L'objectif de cette recherche est de proposer un modèle permettant d'identifier les maladies des feuilles de pommier en utilisant une combinaison d'algorithmes génétiques et d'une méthode de sélection des caractéristiques basée sur la corrélation. Le modèle se compose de quatre étapes principales.

La première étape consiste à acquérir des images de feuilles de pommier, suivie d'un prétraitement et d'une segmentation ponctuelle à l'aide de l'algorithme de croissance de région (RGA). Ensuite, les caractéristiques sont extraites à partir des images ponctuelles.

Dans la troisième étape, la sélection des caractéristiques est réalisée en combinant l'algorithme génétique (GA) et la sélection de caractéristiques basée sur la corrélation (CFS). Cette étape vise à réduire la dimensionnalité de l'espace des caractéristiques en sélectionnant les caractéristiques les plus pertinentes, ce qui améliore la précision de l'identification des maladies des feuilles de pommier.

Enfin, la reconnaissance des maladies est réalisée à l'aide d'un classificateur de machine à vecteurs de support (SVM). Le modèle proposé a été testé sur un ensemble de données comprenant 90 images de maladies, couvrant trois types de maladies des feuilles de pommier : l'oïdium, la mosaïque et la rouille [Alq+22].

Article 3 : Deep learning convolutional neural network for apple leaves disease detection L'étude de recherche visait à classer les feuilles de pommier saines et malades à l'aide d'une architecture de réseau neuronal convolutif (CNN) appelée LeNet. L'ensemble de données utilisé dans l'étude est un sous-ensemble de l'ensemble de données Plant Village et se compose de 1 000 images de feuilles saines et de 1 526 images de feuilles malades, appartenant à trois classes principales : la pourriture noire du pommier, la tavelure du pommier et la rouille du cèdre du pommier. De plus, il existe une classe pour les feuilles de pommier saines. Les chercheurs ont également utilisé des techniques d'augmentation des données pour augmenter la taille de l'ensemble de données et améliorer la précision du modèle. L'architecture LeNet a été appliquée pour classer les images et a atteint une grande précision dans la classification des feuilles de pommier saines et malades. L'étude démontre l'efficacité de l'utilisation des CNN dans la classification des maladies des plantes et le potentiel des techniques d'augmentation des données pour améliorer les performances des modèles [BKA19].

Article 4 : Detecting diseases in apple tree leaves using FPN-ISResNet-Faster RCNN

Le modèle FPN-ISResNet-Faster RCNN est un algorithme d'apprentissage en profondeur spécialement conçu pour détecter les maladies des feuilles de pommier. Il combine plusieurs techniques avancées telles que les réseaux pyramidaux de caractéristiques (FPN), les réseaux résiduels sélectifs améliorés (ISResNet) et les réseaux de neurones convolutifs régionaux plus rapides (Faster RCNN) pour classifier les régions d'intérêt dans une image comme étant saines ou malades. Le processus débute en extrayant les caractéristiques de l'image d'entrée à l'aide du FPN, puis en les affinant avec l'ISResNet. Enfin, le Faster RCNN est utilisé pour classifier les régions d'intérêt en fonction des caractéristiques raffinées.

Pour entraîner et valider le modèle, les chercheurs ont utilisé un ensemble de données d'images de feuilles de pommier provenant de la plateforme Kaggle, qui était associée au concours Plant Pathology 2021 FGVC8. Cet ensemble de données comprend des images de feuilles de pommier saines ainsi que des feuilles atteintes de diverses maladies telles que la rouille du pommier, la rouille du cèdre et la tavelure. Les chercheurs ont prétraité cet ensemble de données et l'ont utilisé pour former le modèle FPN-ISResNet-Faster RCNN.

Les résultats ont démontré que le modèle obtenait une grande précision dans la détection et la classification des feuilles de pommier malades. Cette étude met en évidence le potentiel de l'utilisation d'algorithmes d'apprentissage en profondeur pour identifier les maladies des plantes, et elle peut servir de base pour des recherches futures dans ce domaine [BKA19].

Article 5 : Deep learning for apple diseases : classification and identification

Ce modèle se focalise sur l'utilisation de techniques d'apprentissage en profondeur pour l'identification et la classification précises des maladies de la pomme à partir des feuilles des plantes. La première étape de ce processus consiste à constituer un ensemble de données étiquetées comprenant différents types de maladies de la pomme. Ensuite, une architecture de réseau de neurones convolutifs (CNN) est développée

et entraînée à l'aide de l'apprentissage par transfert afin d'affiner les paramètres du modèle sur l'ensemble de données préparé. Une fois l'apprentissage terminé, le modèle peut être déployé sur les machines des utilisateurs pour classifier les images de feuilles infectées.

Dans le cadre de cette étude, les chercheurs ont préparé un ensemble de données contenant des images de feuilles de pommier affectées par six maladies différentes, notamment la tavelure, l'alternaria, la mosaïque de la pomme, la tache foliaire de Marssonina (MLB), l'oïdium et les feuilles saines. Les images ont été capturées et étiquetées manuellement avec l'aide d'experts du domaine provenant de SKUAST-K.

Grâce à cette approche, le modèle permet d'obtenir une identification et une classification précises des maladies de la pomme à partir des images des feuilles, ce qui peut être utile pour la détection précoce et la prise de décisions en matière de gestion des cultures [KQB21].

Article 6 : Apple leaf disease identification and classification using Resnet models

La méthode proposée pour détecter les maladies des feuilles du pommier comprend plusieurs étapes, telles que le prétraitement de l'image, l'extraction de la zone de la maladie, l'extraction des caractéristiques, la formation et les tests. Au cours de la première étape, l'image est convertie dans l'espace colorimétrique LAB, et l'arrière-plan est supprimé en utilisant une segmentation basée sur un seuil. Dans la deuxième étape, les zones affectées par la maladie sont extraites en éliminant le masque vert. Pour la troisième étape, les caractéristiques de texture et les moments de couleur de la matrice de cooccurrence en niveaux de gris sont calculés afin d'extraire les caractéristiques pertinentes. Ensuite, lors de la quatrième étape, un classificateur SVM est entraîné sur les caractéristiques extraites pour détecter et différencier différents types de maladies des feuilles de pommier. Enfin, le classificateur SVM entraîné est utilisé pour évaluer et classer les images de test en différentes catégories correspondant aux maladies des feuilles de pommier.

Les données utilisées dans cette étude ont été obtenues à partir du site Web Plant Village, qui propose un accès à une collection de plus de 50 000 images de feuilles montrant différentes maladies des plantes. Ces maladies comprennent notamment la maladie des points gris du pommier, la maladie de l'étoile noire, la maladie de la rouille du cèdre, ainsi que des feuilles saines.

Grâce à cette approche, il est possible de détecter et de classifier avec précision les maladies des feuilles du pommier, en utilisant un ensemble de données diversifié et en extrayant des caractéristiques significatives à partir des images des feuilles. Cette méthode offre une solution prometteuse pour la surveillance et la gestion des maladies des cultures de pommiers [LR20].

2.7 Etude comparative

Le tableau suivant montre une étude comparative des méthodes présentée précédemment.

Articles	Techniques	Base de données	Maladies / Classes
[YY+20]	CNN + Fuzzy C-Means clustering	400 images de feuille de pommier	Sain vs malade (non spécifié)
[Alq+22]	SVM + genetic algorithm + correlation-based feature selection	90 images de la maladie de la pomme	Oïdium, mosaïque, rouille
[BKA19]	LeNet CNN + data augmentation	Sous-ensemble de l'ensemble de données Plant Village (1526 malades, 1000 en bonne santé)	Pourriture noire du pommier, tavelure du pommier, rouille du cèdre du pommier, santé
[BKA19]	FPN-ISResNet-Faster RCNN	Images de feuilles de pommier de l'ensemble de données Kaggle (Plant Pathology 2021 FGVC8)	Diverses maladies de la pomme, y compris la rouille, la rouille du cèdre, la tavelure, la santé
[KQB21]	CNN + transfer learning	Images de feuilles de pommier avec six maladies différentes	Tavelure, Alternaria, Mosaïque du pommier, Marssonina Leaf Tache, Oïdium, Santé
[LR20]	SVM + ResNet-18 and ResNet-34 and VGG-16	Images de feuille de pommier	Pourriture noire, rouille du pommier du cèdre, oïdium et tavelure

TABLE 2.1 – Tableau comparatif des méthodes de détection des maladies dans les pommiers

2.8 Conclusion

L'utilisation de l'intelligence artificielle (IA) dans la détection des maladies des pommiers représente une avancée majeure pour la préservation de la santé des cultures. Les différentes approches basées sur l'IA, telles que les réseaux de neurones convolutifs, les algorithmes génétiques et les techniques d'apprentissage en profondeur, permettent une identification et une classification précises des maladies des feuilles de pommier.

Grâce à ces méthodes, il est possible de détecter les signes préliminaires des maladies, souvent invisibles à l'œil nu, et de prendre des mesures préventives ou curatives de manière proactive. Cela aide les arboriculteurs et les agriculteurs à agir rapidement pour prévenir la propagation des maladies, réduire les pertes de récolte et minimiser l'utilisation de produits chimiques nocifs pour l'environnement.

De plus, l'utilisation de l'IA dans la détection des maladies des pommiers permet d'améliorer l'efficacité et la précision du processus de diagnostic. Les modèles d'apprentissage automatique peuvent analyser de grandes quantités de données d'images de feuilles de pommier, identifier les caractéristiques distinctives des maladies et fournir des résultats rapides et fiables. Cela permet de gagner du temps et des ressources, et facilite la prise de décisions éclairées pour la gestion des cultures.

Il convient également de souligner que les techniques développées pour la détection des maladies des pommiers peuvent être adaptées à d'autres types de cultures et de plantes, offrant ainsi des possibilités d'application étendues dans l'agriculture. L'IA ouvre de nouvelles perspectives pour la surveillance continue des cultures, la prévention des épidémies, l'amélioration des rendements et la durabilité environnementale.

CHAPITRE 3

CONCEPTION ET MISE EN ŒUVRE DU SYSTÈME

3.1 Introduction

Ce chapitre se concentre sur la concrétisation de notre système, en mettant en évidence les idées, les concepts et les plans qui faciliteront la transformation de notre vision en réalité. Nous examinerons les différentes étapes du processus de conception et d'exécution, en mettant l'accent sur les décisions prises, les choix technologiques effectués et les mesures prises pour développer un système opérationnel et performant.

3.2 Problématique

Dans le contexte des systèmes de détection des maladies que nous avons présenté dans les chapitres précédents, nous avons constaté que la plupart se concentrent uniquement sur la détection des maladies sur les feuilles des plantes. Cependant, il est important de noter que les maladies qui affectent les branches des arbres peuvent être tout aussi dévastatrices, voire plus. La détection précoce de ces maladies au niveau

des branches peut non seulement sauver la récolte, mais également prévenir la mort des arbres.

C'est pourquoi nous avons pris la décision de nous pencher sur la détection des maladies au niveau des branches des arbres en utilisant les techniques de l'intelligence artificielle. Cette nouvelle approche nous permettra d'élargir notre système de détection et d'offrir une solution plus complète pour la surveillance et la prévention des maladies dans les cultures arboricoles.

3.3 Architecture du système proposé

La méthodologie suggérée a pour objectif de repérer le *Nectria cenabrina* à partir d'image des branches de pommiers.

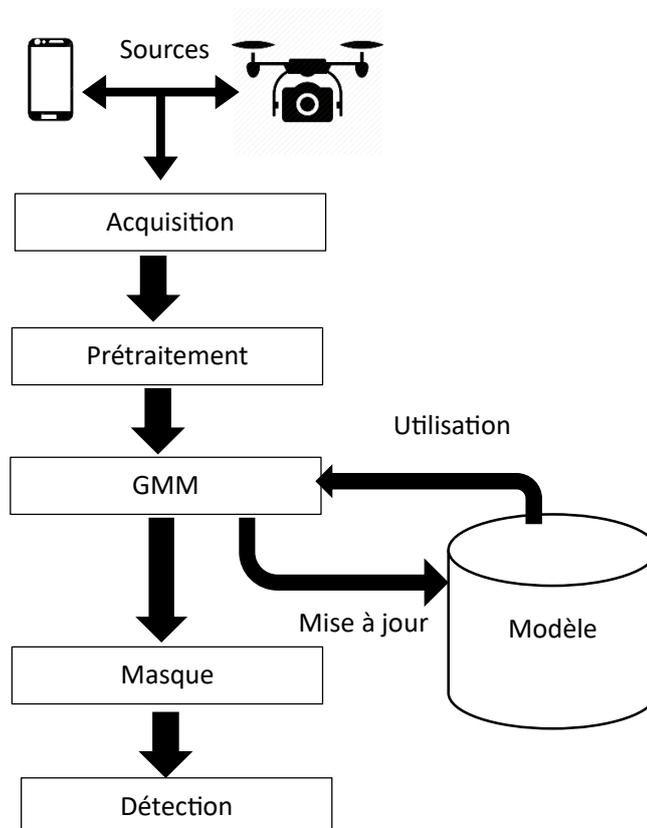


FIGURE 3.1 – Schéma général du processus de détection de la maladie

Le processus se base sur un schéma standard de la RDF à savoir : l'acquisition, le prétraitement, l'extraction des informations pertinentes et la détection.

3.3.1 Acquisition

Le système que nous avons développé est conçu pour traiter des images de type RGB qui sont acquises à l'aide d'un smartphone ou d'une caméra montée sur un drone de surveillance. Ces images sont utilisées comme données d'entrée pour notre système, permettant ainsi d'analyser et de détecter les maladies des branches d'arbres avec une vision globale et détaillée de la situation, que ce soit au sol avec un smartphone ou à distance avec un drone de surveillance.

3.3.2 Prétraitement

Dans la phase de prétraitement, nous apportons des améliorations aux images et extrayons des caractéristiques significatives afin d'améliorer la précision de la détection. Chaque image subit une série d'opérations, telles que celles représentées dans la figure 3.2, qui visent à optimiser sa qualité et à révéler les informations clés nécessaires à la détection des éléments d'intérêt. Ces opérations comprennent des techniques de filtrage, d'amélioration des contrastes et de réduction du bruit, parmi d'autres méthodes avancées. En combinant ces étapes de prétraitement, nous parvenons à obtenir des images prêtes pour l'analyse et la détection, en maximisant la précision et la fiabilité de notre système. Il est à noter que le seuillage et l'opération "et" sont appliqués uniquement dans la phase d'apprentissage du GMM.

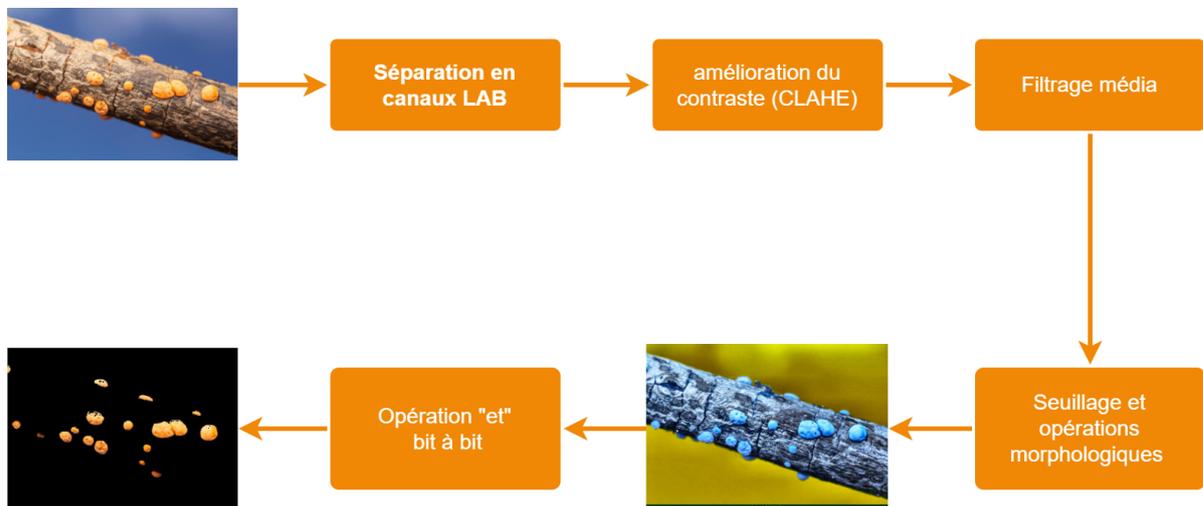


FIGURE 3.2 – Le processus de prétraitement proposé

A Conversion de l'espace de couleur

La conversion de l'espace colorimétrique RVB vers LAB permet de représenter les couleurs de manière plus perceptuellement uniforme. L'espace LAB se compose de trois composantes : L pour la luminosité, A pour la composante verte-magenta et B pour la composante bleu-jaune. Cette transformation permet de séparer la composante de luminosité des informations de couleur, ce qui facilite l'extraction des caractéristiques pertinentes dans le domaine des couleurs. De plus, l'espace LAB est moins sensible aux variations d'éclairage par rapport à l'espace RVB, ce qui améliore la robustesse des méthodes de traitement d'image et de détection des maladies. Ainsi, la conversion de l'espace RVB vers LAB offre une représentation plus adaptée pour les tâches de prétraitement et d'analyse d'images, notamment dans le contexte de la détection des maladies. La conversion des images de l'espace colorimétrique RVB à

l'espace colorimétrique LAB est réalisée en utilisant la formule suivante :

$$\begin{aligned}L &= 0,299 \cdot R + 0,587 \cdot V + 0,114 \cdot B \\a &= (R - L) \cdot 0,5 + 128 \\b &= (B - L) \cdot 0,5 + 128\end{aligned}\tag{3.1}$$

où R , V et B représentent les canaux de couleur rouge, vert et bleu de l'image RVB, et L , a et b représentent les canaux correspondants de l'image LAB.

B Amélioration du contraste

Pour améliorer le contraste local de l'image dans le canal de luminance (L) de l'espace LAB, nous utilisons l'égalisation d'histogramme adaptative limitée au contraste (CLAHE)[MAL18]. Cette technique permet d'augmenter la visibilité des détails dans les régions à faible contraste tout en évitant une surexagération du bruit. L'algorithme CLAHE fonctionne en divisant l'image en petites régions, appelées "tiles", et en égalisant l'histogramme de chaque tile de manière adaptative. Cela signifie que l'égalisation de l'histogramme est effectuée localement sur chaque région, en tenant compte des caractéristiques spécifiques de cette région. En limitant le contraste amélioré à un certain niveau prédéfini, l'algorithme CLAHE préserve les détails structurels de l'image tout en évitant une amplification excessive du bruit. Le pseudo-code suivant donne une idée générale de la procédure CLAHE :

1. Diviser l'image en tiles de taille fixe.
2. Calculer l'histogramme de chaque tile.
3. Répartir l'histogramme en utilisant une fonction de transfert adaptative pour chaque tile.
4. Réaliser une interpolation bilinéaire entre les valeurs des tiles voisines pour éviter les artefacts de bord.

5. Réassembler les tiles égalisées pour obtenir l'image finale avec un contraste amélioré.

L'application du CLAHE au canal L de l'espace LAB permet d'améliorer la qualité visuelle de l'image en mettant en évidence les détails et en augmentant la perception des variations locales de luminance. Cette étape de prétraitement contribue à améliorer la précision de la détection des maladies en révélant les caractéristiques importantes dans les régions d'intérêt.

C Réduction du bruit

Afin de réduire le bruit dans les canaux "a" et "b" de l'image LAB, nous appliquons une technique de filtrage médian. Cette opération permet de remplacer chaque valeur de pixel par la médiane des valeurs des pixels voisins, ce qui contribue à atténuer les fluctuations indésirables. Le filtrage médian est un processus non linéaire qui prend en compte les valeurs des pixels environnants pour estimer la valeur du pixel en cours de traitement. Cela permet d'éliminer les valeurs aberrantes et les variations brusques, tout en préservant les contours et les détails importants de l'image. Le pseudo-code suivant présente les étapes générales du filtrage médian :

1. Parcourir chaque pixel de l'image dans les canaux "a" et "b".
2. Pour chaque pixel, extraire le voisinage autour du pixel en utilisant une fenêtre de taille définie.
3. Trier les valeurs des pixels voisins en ordre croissant.
4. Sélectionner la valeur médiane parmi les valeurs triées.
5. Remplacer la valeur du pixel en cours de traitement par la valeur médiane obtenue.
6. Répéter les étapes précédentes pour tous les pixels de l'image.

L'application du filtrage médian aux canaux "a" et "b" de l'image LAB permet d'atténuer le bruit tout en préservant les informations utiles liées à la couleur et à la texture. Cette étape de prétraitement améliore la qualité de l'image en réduisant les variations indésirables, ce qui facilite la détection précise des maladies sur les branches de l'arbre.

D Seuillage et opérations morphologiques

Le processus de seuillage est appliqué sur le canal B (canal bleu-jaune) de l'image LAB pour créer une image binaire, où les pixels ayant des valeurs supérieures à un seuil spécifique sont considérés comme des zones bleues potentiellement représentant la présence du *Nectria cenabrina*. Le pseudo-code suivant illustre les étapes générales du seuillage :

1. Parcourir chaque pixel de l'image dans le canal B.
2. Comparer la valeur du pixel avec un seuil prédéfini.
3. Si la valeur du pixel est supérieure au seuil, assigner la valeur 1 (zone bleue) à ce pixel dans l'image binaire ; sinon, assigner la valeur 0 (zone non bleue).
4. Répéter les étapes précédentes pour tous les pixels de l'image.

Après la création de l'image binaire, des opérations morphologiques sont appliquées pour améliorer la qualité de la détection. La fermeture morphologique est utilisée pour connecter les zones bleues qui pourraient ne pas être complètement connectées, créant ainsi des formes plus solides et cohérentes. En revanche, l'ouverture morphologique est appliquée pour éliminer les petits trous présents à l'intérieur des zones bleues, permettant ainsi d'obtenir des formes plus lisses et uniformes. Le pseudo-code ci-dessous présente les étapes générales des opérations morphologiques :

1. Appliquer la fermeture morphologique à l'image binaire.
2. Utiliser un élément structurant (par exemple, un noyau de forme circulaire) pour élargir les zones bleues et les connecter.
3. Répéter cette opération pour un certain nombre d'itérations prédéfini.

4. Appliquer l'ouverture morphologique à l'image résultante de la fermeture.
5. Utiliser un élément structurant pour supprimer les petits trous à l'intérieur des zones bleues.
6. Répéter cette opération pour un certain nombre d'itérations prédéfini.

E Opération "et" bit à bit

Pour mettre en évidence les régions d'intérêt, une opération "et" bit à bit est appliquée entre l'image traitée et l'image binaire obtenue précédemment. Cette opération permet de conserver uniquement les pixels qui sont bleus dans l'image binaire, masquant ainsi le reste de l'image d'origine. Le pseudo-code suivant décrit les étapes générales de l'opération "et" bit à bit :

1. Parcourir chaque pixel de l'image traitée et de l'image binaire.
2. Effectuer une opération "et" bit à bit entre les valeurs des pixels correspondants dans les deux images.
3. Assigner la valeur résultante à ce pixel dans une nouvelle image qui servira de masque.
4. Répéter les étapes précédentes pour tous les pixels des deux images.

L'opération "et" bit à bit crée une image masquée où seules les régions correspondant aux pixels bleus dans l'image binaire sont visibles. Cela permet de mettre en évidence les zones potentielles de présence du *Nectria cenabrina* dans l'image d'origine. En se concentrant sur ces régions d'intérêt, il est plus facile de procéder à une analyse plus précise et à une détection plus fiable de la maladie sur les branches de l'arbre.



(A) image avant le traitement



(B) image après le Seuillage



(C) image après le traitement

FIGURE 3.3 – pretraitement étape

3.3.3 Gaussian Mixture Model

Stauffer et Grimson [SG99] ont été les précurseurs de l'utilisation des mélanges de gaussiennes (GMM) pour la modélisation de l'arrière-plan. Les GMM se sont révélées efficaces pour gérer les variations de luminosité, les mouvements répétitifs et les situations encombrées. Cette approche a démontré sa capacité à fournir une représentation robuste de l'arrière-plan en adaptant dynamiquement les différentes gaussiennes pour prendre en compte les changements et les perturbations. L'introduction des GMM par Stauffer et Grimson a marqué une avancée majeure dans la modélisation de l'arrière-plan et a ouvert la voie à de nombreuses applications dans le domaine de la vision par ordinateur.

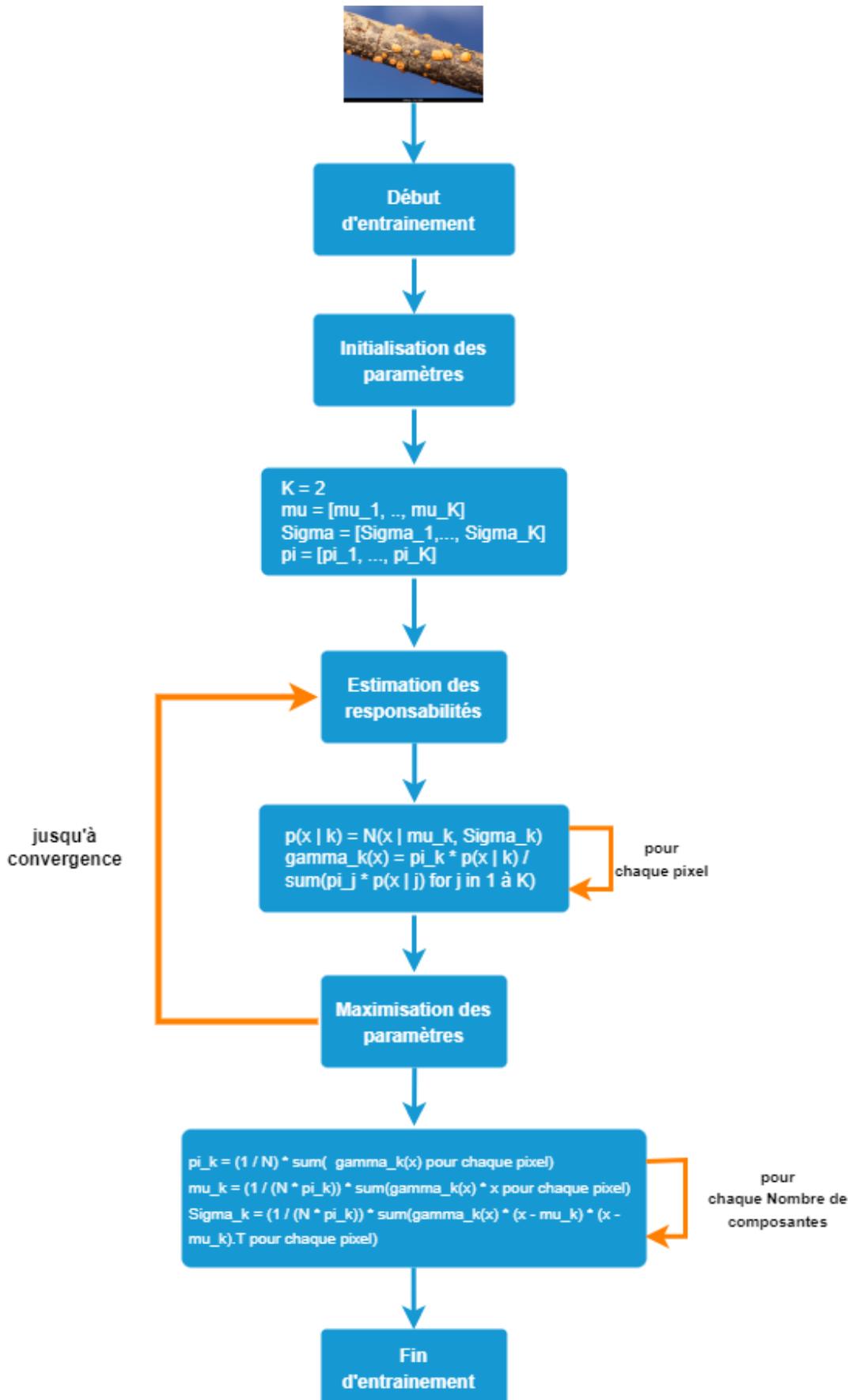


FIGURE 3.4 – Les étapes du GMM

La motivation sur laquelle est basée la proposition de ce modèle est qu'un modèle unimodal de la maladie ne pouvait gérer le bruit d'acquisition d'image, le changement de lumière et l'existence de multitude de surfaces pour un pixel particulier dans le même temps.

Une GMM est un modèle statistique qui suppose que les données proviennent d'une somme pondérée de plusieurs distributions de gaussiennes. Si les K distributions de gaussiennes sont utilisés pour décrire l'historique d'un pixel, l'observation d'un pixel donné sera forcément dans l'un des K états à la fois. K détermine la multimodalité de la maladie et la sélection de K est généralement basé sur la mémoire disponible et de la puissance de calcul.

A Initialisation

Tout d'abord, chaque pixel est caractérisé par son intensité dans l'espace de couleurs LAB. Ensuite, la probabilité d'observer la valeur du pixel courant X_t est donnée par l'équation suivante dans le cas multidimensionnel.

$$P(X_t) = \sum_{i=1}^k \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (3.2)$$

Où : k est le nombre de gaussiennes associées à chaque pixel, $\omega_{i,t}$ est le poids calculé, $\mu_{i,t}$ est la moyenne et $\Sigma_{i,t}$ est la matrice de covariance qui sont respectivement évalués pour la $i^{\text{ème}}$ gaussienne à l'instant t . η est la fonction de densité de probabilité gaussienne :

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (3.3)$$

Où D exprime la dimension du vecteur X .

Au début du processus de détection de la maladie, pour chaque pixel, les distributions des K gaussiennes sont initialisées avec une moyenne prédéfinie, une variance élevée et un faible poids. Pour obtenir un modèle stable, nous avons utilisé plusieurs images pour l'apprentissage des mixtures de gaussiennes.

B Mise à jour des paramètres

Il existe plusieurs techniques disponibles pour estimer les paramètres d'une GMM. La méthode la plus populaire et la plus utilisée pour l'estimation est le maximum de vraisemblance (ML). Le but de l'estimation ML est de trouver les paramètres du modèle qui maximisent la vraisemblance du GMM sur un ensemble de données d'apprentissage. Pour une séquence de T vecteurs d'apprentissage de $X = x_1, \dots, x_T$, si on suppose qu'il existe une indépendance entre les vecteurs, alors la vraisemblance du GMM peut être écrite selon l'équation 3.4 suivante.

$$P(X|\lambda) = \prod_{t=1}^T P(X_t|\lambda) \quad (3.4)$$

Malheureusement, cette expression est une fonction non linéaire des paramètres λ et la maximisation directe est impossible. Cependant, les estimations des paramètres ML peuvent être obtenues en utilisant de manière itérative un cas particulier de l'algorithme espérance-maximisation (EM).

L'idée de base de l'algorithme EM est de commencer par un modèle initial λ pour estimer un nouveau modèle $\hat{\lambda}$, tel que $P(X|\hat{\lambda}) \geq P(X|\lambda)$. Le nouveau modèle devient alors le modèle initial pour la prochaine itération et le processus est répété jusqu'à ce qu'un seuil de convergence soit atteint. Le modèle initial est généralement obtenu en utilisant une forme binaire de l'histogramme. Pour chaque itération d'EM, la ré-estimation des équations 3.5, 3.6 et 3.7 est utilisée pour garantir une augmentation monotone de la valeur de vraisemblance du modèle.

$$\bar{\omega}_i = \frac{1}{T} \sum_{t=1}^T \hat{P}(i|X_t, \lambda) \quad (3.5)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T \hat{P}(i|X_t, \lambda) X_t}{\sum_{t=1}^T \hat{P}(i|X_t, \lambda)} \quad (3.6)$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T \hat{P}(i|X_t, \lambda) X_t^2}{\sum_{t=1}^T \hat{P}(i|X_t, \lambda)} - \bar{\mu}_i^2 \quad (3.7)$$

Ou $\bar{\omega}_i$, $\bar{\mu}_i$ et $\bar{\sigma}_i^2$ représente respectivement les nouvelles valeurs du poids, de la moyenne et de la variance. La probabilité a posteriori pour le composant i est donnée par l'équation suivante :

$$\hat{P}(i|X_t, \lambda) = \frac{\omega_i \eta(X_t, \mu_{i,t}, \Sigma_{i,t})}{\sum_{k=1}^M \omega_k \eta(X_t, \mu_{k,t}, \Sigma_{k,t})} \quad (3.8)$$

Malgré la réputation de l'algorithme ML ce dernier occupe énormément d'espace mémoire et nécessite une puissance de calcul non négligeable ce qui rend son utilisation inappropriée dans un environnement en temps réel ou d'un système embarqué. En plus, EM converge vers un maximum local ; mais nous n'avons aucune garantie que ce maximum soit « bon », c-a-d soit proche en valeur du maximum global. Pour se rapprocher de ce but, on doit encore alourdir le calcul. Enfin, EM présente le risque de trouver des solutions « dégénérées » présentant au moins une composante ajustée sur un seul individu ; or la vraisemblance d'une telle Gaussienne peut être infinie (i.e. on se rapproche d'un pic de Dirac), mettant ainsi à mal notre algorithme.

Pour une éventuelle application en temps réel, nous avons utilisé l'algorithme d'approximation K-Means pour la mise à jour du modèle. Chaque valeur d'un nouveau pixel est vérifiée, à travers les k distributions de gaussiennes, pour déterminer si la distance entre le pixel et chacune des gaussiennes est inférieure à l'écart type de distribution selon l'équation suivante :

$$\frac{|P_t - \mu_i|}{\sigma_i} < 2.5 \quad (3.9)$$

Si aucune distribution ne vérifie l'équation 3.9, alors le pixel est associé à la branche et non pas à la maladie. Les paramètres, de la gaussienne la moins probable, sont remplacés par la moyenne, la variance et le poids du pixel courant suivant les équations 3.10, 3.11 et 3.12

$$\mu_k = X_t \quad (3.10)$$

$$\sigma_k^2 = \text{une grande variance initiale} \quad (3.11)$$

$$\omega_k = \text{faible poids} \quad (3.12)$$

Les poids de toutes les gaussiennes seront mis à jour selon l'équation suivante :

$$\omega_{k,t} = (1 - \alpha) \cdot \omega_{k,t-1} + \alpha M_{k,t} \quad (3.13)$$

Où α est le coefficient d'apprentissage qui détermine la vitesse d'adaptation du modèle.

$M_{k,t}$ est égal à 1 pour la gaussienne qui correspond à la maladie et égal à 0 pour les autres. Après la mise à jour des poids, une étape de normalisation est effectuée pour garantir que la somme des poids est toujours égale à 1.

Les autres paramètres des distributions ayant vérifié l'équation 3.9 sont mis à jour selon les équations 3.14, 3.15 et 3.16.

$$\mu_{k,t} = (1 - \varphi_k) \cdot \mu_{k,t-1} + \varphi_k \cdot P_t \quad (3.14)$$

$$\sigma_{k,t}^2 = (1 - \varphi_k) \cdot \sigma_{k,t-1}^2 + \varphi_k (P_t - \mu_{k,t})^T (P_t - \mu_{k,t}) \quad (3.15)$$

avec

$$\varphi_t = \alpha \cdot \eta (P_t | \mu_k, \sigma_k) \quad (3.16)$$

3.3.4 Création du masque

Pour décider si X_{tj} représente une maladie, les distributions sont ordonnées selon la valeur $w_{k,t} / \sigma_{k,t}$. Cet ordre utilise l'hypothèse selon laquelle un pixel de la maladie correspond à un poids élevé avec une variance faible pour la raison que la couleur de la maladie est pratiquement constante. Les premières B distributions qui vérifient l'équation 3.17 sont sélectionnées pour représenter la maladie.

$$B = \arg \min_b \sum_{k=1}^b \omega_{k,t} > T \quad (3.17)$$

Le seuil T représente la portion minimum du poids total attribué au modèle de la maladie. Si une petite valeur T est choisie, la maladie devient unimodal. L'utilisation

d'un seuil T unique pour le modèle engendre une mauvaise classification surtout lorsque la scène contient à la fois des zones dynamique et statique. Un seuil plus élevé permet d'avoir une meilleure classification dans un environnement dynamique, mais génère des détections incorrectes de la maladie.

3.3.5 Detection

Les masques binaires résultants pour chaque composante sont remodelés pour correspondre aux dimensions de l'image originale. Ces masques binaires sont ensuite appliqués à l'image originale à l'aide d'une opération de ET logique, ce qui donne une image masquée mettant en évidence les régions infectées par *Nectria cenabrina*. L'image masquée peut être visualisée et analysée pour évaluer les résultats de détection.

3.4 Conclusion

Nous avons proposé une approche qui permet de détecter la présence de la maladie *Nectria cenabrin* dans les branches d'arbre de pommiers. Nous avons commencé par une étape de prétraitement qui permet d'améliorer l'image en termes de qualité et de précision afin de faciliter la tâche à notre classifieur. La classification est entièrement basée sur les mélanges de gaussiennes qui apprennent la distribution de probabilité de la couleur de la maladie, ce qui génère un masque de segmentation. Ce dernier est utilisé pour localiser les zones malades de la branche.

CHAPITRE 4

IMPLÉMENTATION, TESTS ET RÉSULTATS

4.1 Introduction

Ce chapitre se focalise sur la mise en œuvre concrète de notre système, mettant en lumière les étapes clés, les tests effectués et les résultats obtenus. Après avoir défini la conception et les concepts fondamentaux dans les chapitres précédents, il est temps de concrétiser notre vision en transformant nos idées et nos plans en une réalité fonctionnelle.

Nous entamerons en décrivant minutieusement le processus d'implémentation, mettant en évidence les choix technologiques et les outils sélectionnés pour développer notre système, ainsi que les bases utilisées et les valeurs des paramètres choisis. Nous aborderons les aspects techniques, les défis rencontrés et les solutions mises en place pour surmonter ces obstacles.

4.2 Paramétrage

L'étape de paramétrage du modèle revêt une grande importance pour garantir de bonnes performances, quel que soit le type d'application d'intelligence artificielle.

Afin d'obtenir les meilleurs résultats possibles, nous avons effectué de nombreux tests et avons fixé les paramètres de manière empirique. Après une analyse minutieuse, nous avons déterminé les valeurs suivantes pour optimiser notre système :

- **le nombre de composants du mélange : 2**
- **le type de covariance : *full***
- **Le seuil de convergence : $1e-3$**
- **Régularisation non négative ajoutée à la diagonale de covariance : $1e-6$**
- **Le nombre d'itérations EM à effectuer : 100**
- **Le nombre d'initialisations à effectuer : 1**
- **La méthode utilisée pour initialiser les poids, les moyennes et les précisions : *kmeans***
- **Nombre d'itérations effectuées avant la prochaine impression : 10**

Pour les paramètres des autres méthodes, nous avons utilisé les valeurs par défaut proposer par les API de Python.

4.3 Outils Matériels

Les expérimentations ont été conduites sur un ordinateur portable fonctionnant sous Windows 10. L'ordinateur en question est équipé d'un processeur i7 7700HQ cadencé à 2,80 GHz et de 8 Go de mémoire RAM DDR4. Pour l'implémentation des algorithmes, nous avons utilisé le langage Python au sein de l'environnement de développement Jupiter et Google Colab.

4.4 Outils Logiciels

4.4.1 Environnement de développement

Jupyter : Logiciels libres, normes ouvertes et services Web pour le calcul collaboratif de tous les langages de programmation. Le Jupyter Notebook est la première application Web conçue pour créer et partager des documents informatiques. Il offre une expérience simple, rationalisée et axée sur les documents.[Jup23]

4.4.2 Langage de programmation

Python : Python est un langage de programmation de haut niveau orienté objet, interprétable et doté d'une sémantique flexible. Ses structures de données intégrées, son encodage dynamique et sa liaison en font un langage souhaitable pour le développement rapide d'applications et la création de scripts. Sa syntaxe simple et facile à apprendre met l'accent sur la clarté et permet d'économiser sur les dépenses de maintenance. Il propose des modules et des packages, favorisant ainsi la modularité du programme et la réutilisation du code. L'interpréteur Python et la bibliothèque commune sont accessibles sous forme source ou binaire et peuvent être distribués sans restriction.[The]

4.4.3 Description des bibliothèques utilisées

Scikit-learn : Scikit-learn est une bibliothèque libre Python destinée à l'apprentissage automatique. Elle est développée par de nombreux contributeurs, notamment dans le monde académique par des instituts français d'enseignement supérieur et de recherche comme l'Inria et Télécom ParisTech . [Bis19]

OpenCV : OpenCV est une bibliothèque de vision par ordinateur qui est disponible sous la forme d'une bibliothèque open source et est accessible pour tout le monde. La

bibliothèque utilise les langages C et C++ et est compatible avec Linux, Windows et Mac OS X. [BK08]

Numpy : Ceci est une bibliothèque de langues Programmation Python, ajoutant la prise en charge des grands tableaux multidimensionnels et Matrices et une multitude de fonctions mathématiques avancées Opérez sur ces tables. L'ancêtre de NumPy, Numeric, a été écrit par Jim Hugunin Contributions de plusieurs autres développeurs. En 2005, Travis Oliphant Combinant les fonctionnalités de Numarray concurrentes de Numeric, changements importants. NumPy est un logiciel open source, il en existe de nombreux Donateur . [Num23]

Mathplotlib : Mathplotlib est une création de John D. Hunter en 2003 qui a été publiée le 8 mai 2021 et a une version actuelle de 3.7.1. Cette bibliothèque est principalement écrite en Python, tandis que le reste est écrit en objectif C et JavaScript, ce qui la rend compatible avec la plate-forme. [Joh17]

4.5 La base de données

Nous avons mené de multiples recherches afin de trouver une base de données contenant un nombre suffisant d'images de branches malades en général, et plus spécifiquement de branches de pommiers. Malheureusement, ces recherches se sont soldées par un échec, car aucune base de données correspondante n'était disponible. Par conséquent, nous avons pris la décision de créer notre propre base de test.

Nous avons pu collecter environ 650 images à partir d'internet. Parmi celles-ci, nous avons sélectionné 500 images présentant un seul type de maladie, couvrant différents cas, afin de constituer notre base de données.

Les variations entre les cas ont été basées sur la couleur et l'arrière-plan. Les couleurs obtenues comprennent des nuances d'orange sombre, d'orange clair, de rouge sombre, de rouge clair et de jaune.

Quant à l'arrière-plan, certaines images comportent un arrière-plan vide ou d'une seule couleur, tandis que d'autres présentent un arrière-plan avec des arbres et des feuilles.

Le tableaux suivants récapitulent toutes les différences entre les images dans notre base de données :

Type Image	Nombre total	Entraînement	Tests
Arrière-plan contenant des arbres et des feuilles	116	80	36
Arrière-plan vide ou autre	384	264	120
Totale	500	344	156

TABLE 4.1 – Nombre d'images par variante d'arrière plan

Type Image	Nombre total	Entraînement	Tests
Orange sombre	230	170	60
Orange claire	100	60	40
Rouge sombre	60	40	20
Rouge claire	80	50	30
Jaune	30	24	6
Totale	500	344	156

TABLE 4.2 – Nombre d'images par variante de maladie

Grâce à cette base de données que nous avons créée, nous serons en mesure de mener nos expériences et d'évaluer les performances de notre système de détection des maladies des branches de pommiers.

4.6 Les Métriques

Recall : En classification et en apprentissage automatique, le Recall est une mesure de performance utilisée pour évaluer la capacité d'un modèle à identifier tous les

exemples positifs parmi tous les exemples réels positifs.

Le calcul est obtenu en divisant le nombre de vrais positifs par la somme des vrais et faux négatifs. En d'autres termes, il calcule le taux de résultats positifs approprié pour chaque exemple réel positif.

$$\text{Recall} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux ngatifs}} \quad (4.1)$$

Précision : La précision est une mesure de performance utilisée en classification et en apprentissage automatique pour évaluer la capacité d'un modèle à identifier correctement les exemples positifs parmi toutes les prédictions positifs. La somme des vrais et des faux positifs est divisée par le nombre de vrais positifs pour calculer la précision. En d'autres termes, elle mesure le taux de résultats positifs exacts parmi toutes les prédictions positives. La précision élevée indique que le modèle est capable de produire un faible taux de faux positifs, ce qui est généralement souhaitable. Cependant, une précision élevée peut également entraîner une augmentation des faux négatifs, c'est-à-dire les cas où le modèle prédit une classe négative alors que l'exemple réel est positif.

$$\text{Precision} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux positifs}} \quad (4.2)$$

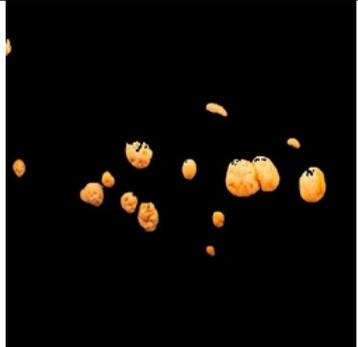
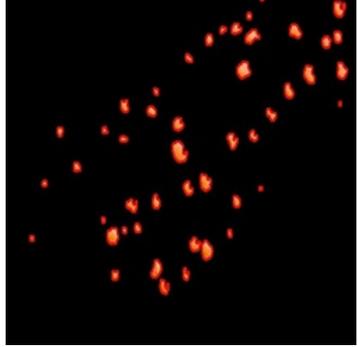
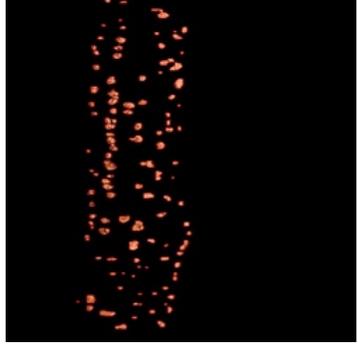
F1-Score : Le score F1 est une moyenne harmonique de la précision et du rappel. Il est utilisé pour prendre en compte à la fois la précision et le rappel, mettant davantage l'accent sur les classes inégales.

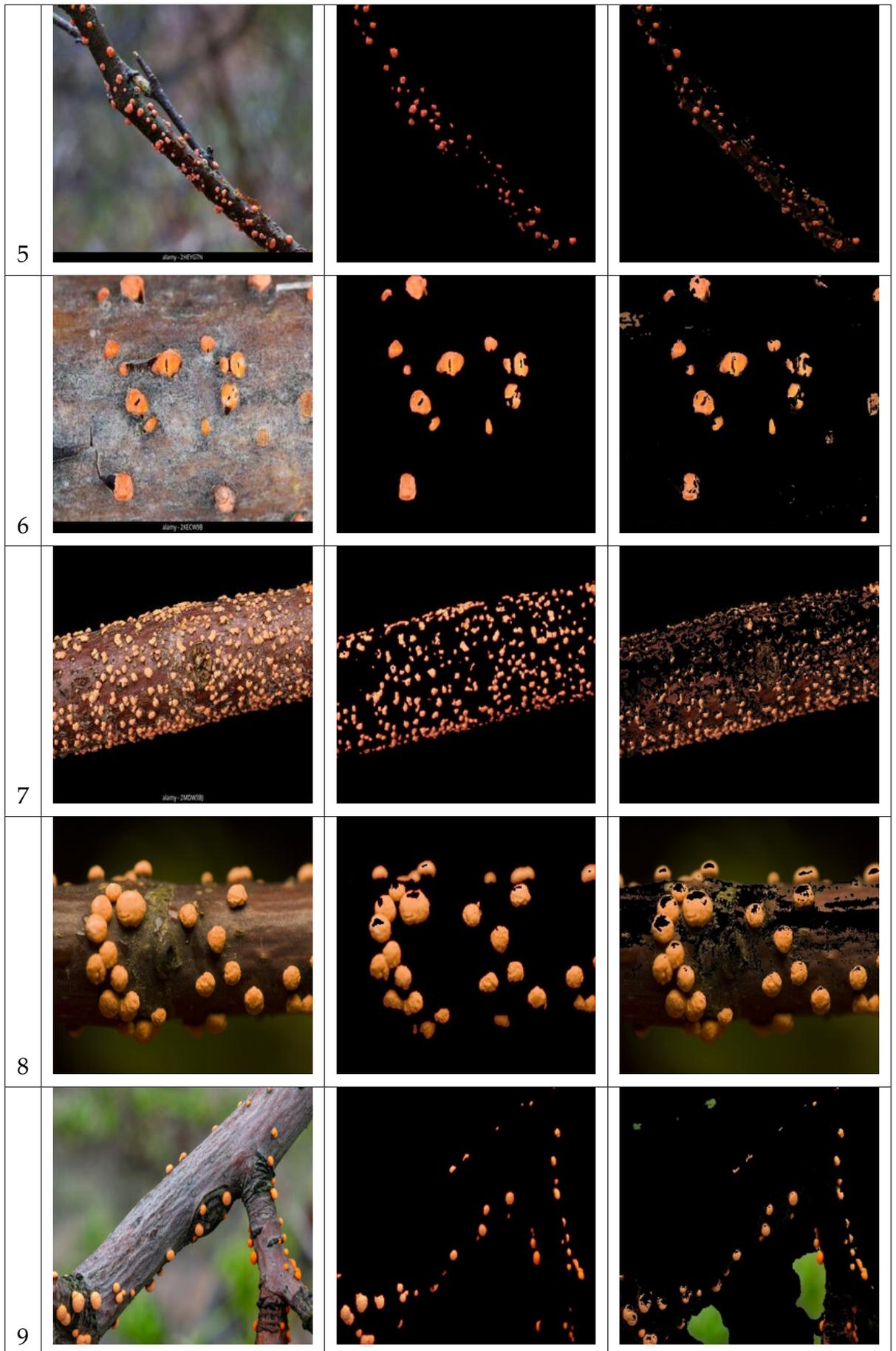
Une valeur de 1 sur le F1-score indique une classification parfaite, tandis qu'une valeur de 0 indique une classification mauvaise. Par conséquent, un score F1 plus élevé signifie que le modèle de classification fonctionne mieux en termes de précision et de rappel.

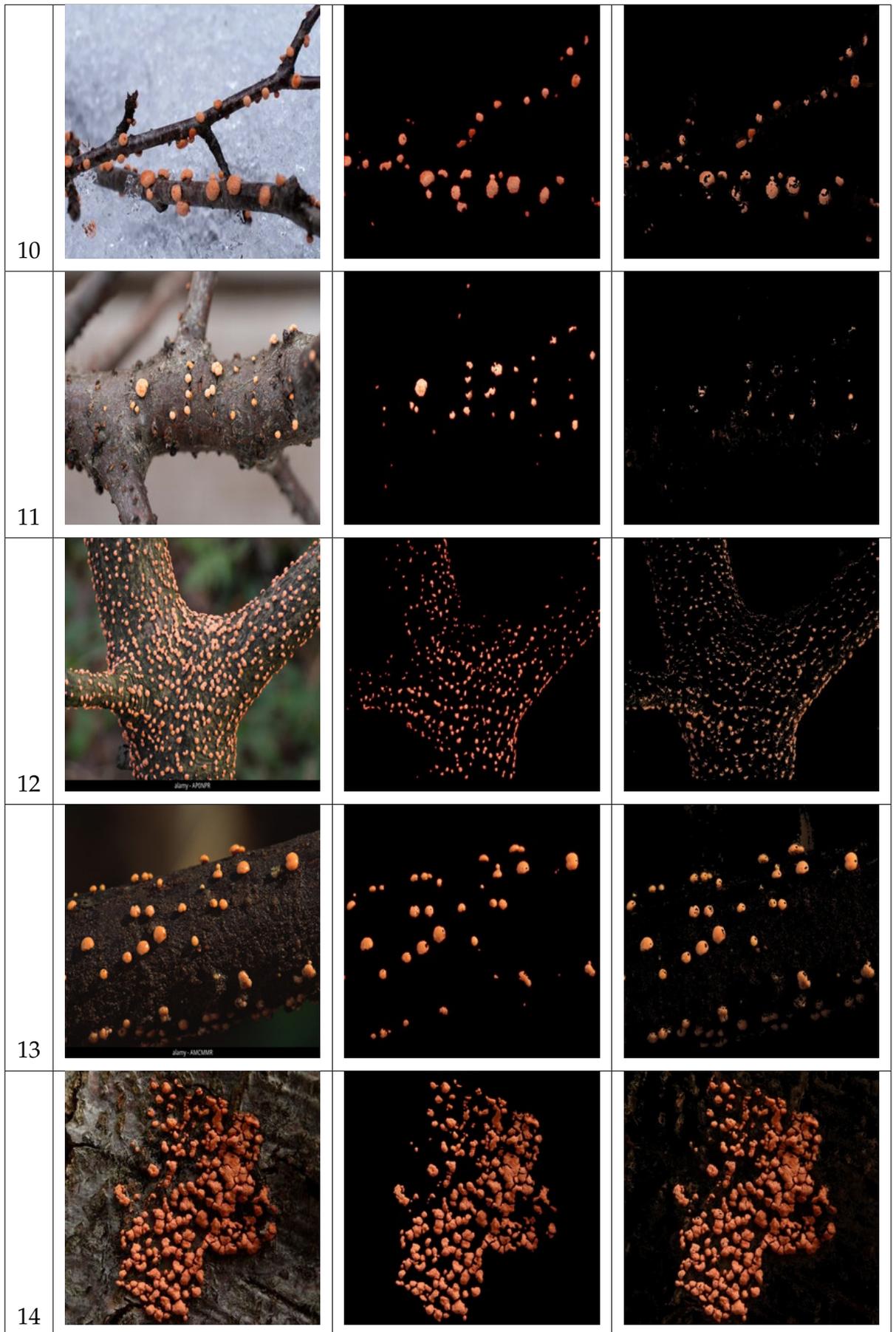
$$F1\text{-score} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \quad (4.3)$$

4.7 Résultats Qualitatifs

Cette section met en évidence les aspects qualitatifs et visuels des résultats obtenus par notre système de détection des maladies des branches de pommiers.

#	Image originale	vérité terrain	Detection
1			
2			
3			
4			





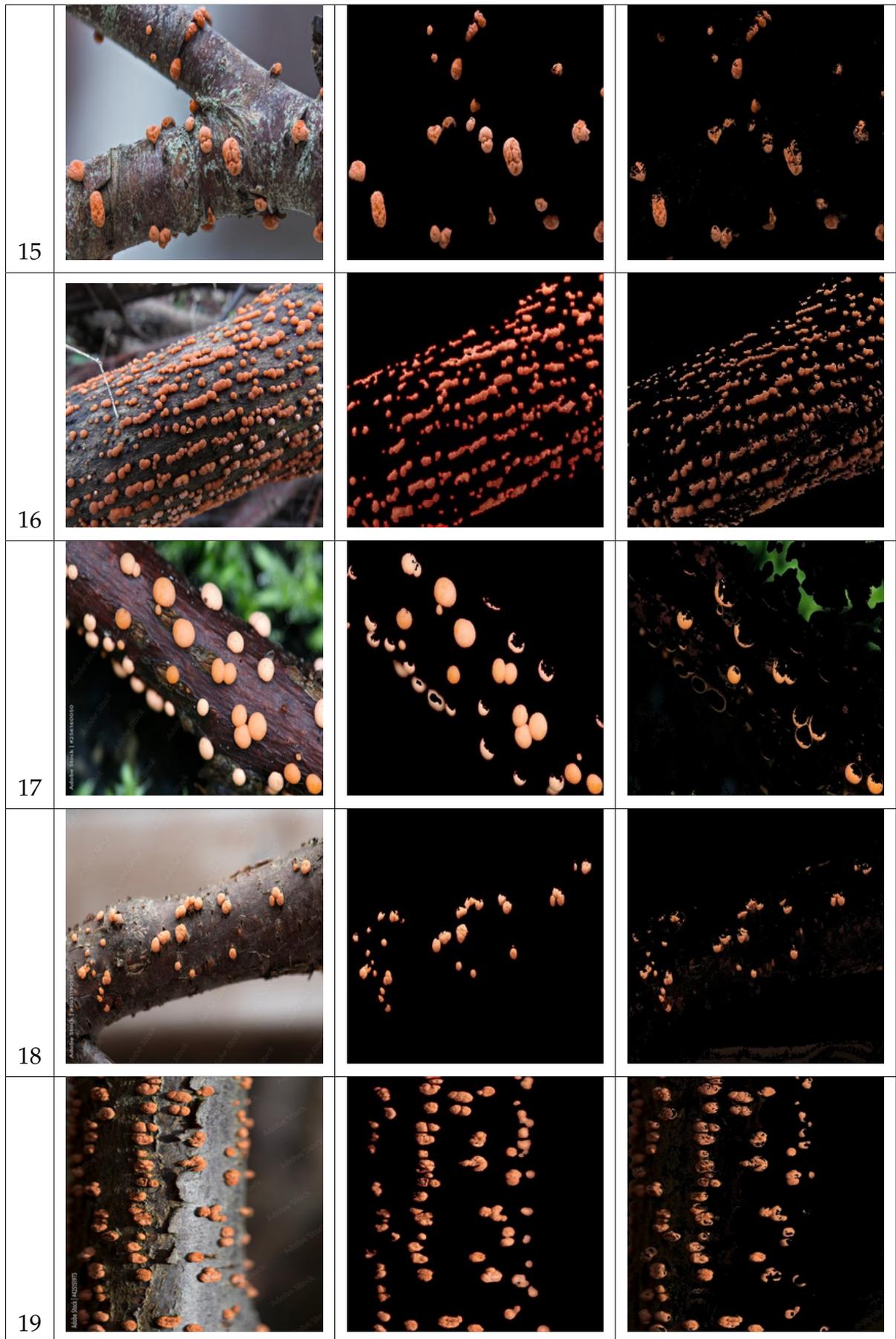


TABLE 4.3 – Quelques résultats de la detection de la maladie Nectria canabrina dans les branche des pommiers

Effectivement, le système a réussi à détecter la présence de la maladie dans toutes les images de test, ce qui témoigne de la robustesse de la méthode utilisée. Cela est dû à sa capacité à mémoriser la distribution de probabilité des pixels représentant la maladie. Cependant, nous avons également remarqué que le système a détecté des pixels en dehors de la zone malade, notamment dans les images présentant une végétation abondante.

Cette observation peut être attribuée à une réduction de l'écart entre les couleurs lors de la conversion de l'image du modèle RVB vers le modèle HSV. Lorsque nous recherchons la couleur orange en utilisant la représentation RVB, nous ajustons les valeurs des canaux Rouge et Vert, tandis que le canal Bleu est fixé à zéro. Une combinaison courante de valeurs RVB pour obtenir la couleur orange est Rouge = 255, Vert = 165, Bleu = 0. Il est donc vrai que la couleur verte est présente dans la couleur orange.

Cela peut entraîner une certaine confusion pour le système de détection, car certaines parties de la végétation peuvent avoir des valeurs de couleur similaires à celles de la couleur orange. Cela peut expliquer la détection de pixels en dehors de la zone malade, mais dans des régions avec une certaine verdure. Il convient de prendre en compte cette limitation lors de l'évaluation et de l'interprétation des résultats du système de détection.

4.8 Résultats Quantitatifs

Dans cette section, nous présenterons les résultats en terme de rappel (recall), précision et le F-score pour évaluer les performances de notre système de détection des maladies des branches de pommiers. Ces statistiques nous permettront de quantifier l'efficacité de notre approche et d'identifier les aspects qui nécessitent d'éventuelles améliorations.

#	F1-score	Recall	Précision
1	0.7990421455938698	0.7845387002727359	0.8140919293451742
2	0.7801719325667077	0.8515720204728248	0.7198187062216729
3	0.650347705146036	0.5089246843709185	0.9006163328197226
4	0.9366577906882918	0.893679267072138	0.9839789777061965
5	0.7502691065662002	0.6619183285849952	0.8658385093167702
6	0.8581730769230769	0.8521068576149821	0.8643262873638141
7	0.6115045526810232	0.46205089002948563	0.9038666951506088
8	0.7705454545454545	0.6440729483282674	0.9588235294117647
9	0.7088399605652317	0.7060556464811784	0.7116463213460904
10	0.7140992167101827	0.5723753051970701	0.9491035280508965
11	0.2976092333058532	0.1791563275434243	0.878345498783455
12	0.7486553578816715	0.6263412945655936	0.9303341902313624
13	0.6421428571428572	0.5461725394896719	0.7790294627383015
14	0.6818613485280152	0.7223340040241448	0.64568345323741
15	0.7759965022225461	0.8534220227600577	0.7114510956707643
16	0.37337461300309593	0.23399755524942276	0.923359620243473
17	0.6354403025391681	0.4762714609653385	0.9543979227523531
18	0.7845248685175648	0.6551202856621576	0.9776351614157915
19	0.7160828333655893	0.6441504178272981	0.8061002178649237
20	0.5276715636949084	0.36457174638487205	0.954843408594319
21	0.39857489387507583	0.6013266239707228	0.29807256235827667
22	0.673263272826532	0.5674680777637179	0.8275457138064083
23	0.6941795168909884	0.6650725875320239	0.7259507829977628
24	0.7120726495726495	0.6835897435897436	0.7430323299888517
25	0.5902340800402719	0.4262086513994911	0.95949263502455
26	0.7500503524672709	0.6248322147651006	0.9380352644836272
27	0.9701637686233321	0.9922024188415022	0.9490828830200168
28	0.9593978170869402	0.947600071373342	0.9714930333241867
29	0.7796239411709952	0.6802160318362707	0.9130600675896653
30	0.9101209097827007	0.8894397620228062	0.931786703601108
31	0.5738491017498129	0.408998266897747	0.9613178648243499
32	0.8107698403128283	0.7587129638585292	0.8704964119257231
33	0.2702587128501176	0.18246559522663844	0.5208791208791209
34	0.6090504451038575	0.48579881656804735	0.81610337972167
35	0.8551427032848681	0.9962358845671268	0.7490566037735849
36	0.40912121616215497	0.26075131735509094	0.9492574257425742
37	0.1552839788335141	0.4587956989247312	0.09345786204298344
38	0.9133431889281053	0.9014209538205004	0.9255850190772111
39	0.7175024247031377	0.6401472951941967	0.8161222996232348
40	0.922270260325637	0.9257617728531856	0.9188049853372434
41	0.5307411907654921	0.5245796637309847	0.5370491803278689
42	0.7611786174020058	0.7383572545898625	0.785455713201378
43	0.6349738633001282	0.5725211204979991	0.712720026569246
44	0.5491932715413662	0.8215055972065318	0.41246841643892124
Moyenne	0.67920979	0.641056639	0.801898086

TABLE 4.4 – Résultats en F1-score, Recall et Précision dans quelques images de tests

4.9 Conclusion

Au cours de ce chapitre, nous avons exploré en détail l'implémentation de notre système, les outils utilisés ainsi que les paramètres qui ont été fixés. Nous avons soumis notre système à une base de tests minutieusement collectée, image par image. Cependant, l'une des tâches les plus exigeantes a été la création du référentiel de vérité terrain pour chaque image, ce qui a demandé une concentration soutenue, ainsi qu'un investissement considérable en temps et en effort.

Les résultats obtenus ont été globalement satisfaisants, témoignant de l'efficacité de notre système de détection. Nous avons réussi à identifier avec précision les zones malades dans les images de test, démontrant ainsi la capacité de notre approche à reconnaître les symptômes des maladies des branches de pommiers. Cette réalisation ouvre également la voie à de nouvelles perspectives pour améliorer notre système de détection et pour l'adapter à d'autres types de maladies.

CONCLUSION GÉNÉRALE

Ce mémoire a exploré en profondeur l'utilisation de l'intelligence artificielle (IA) pour la détection des maladies sur les branches des pommiers. Nous avons présenté un système intelligent basé sur l'IA qui utilise des techniques d'apprentissage automatique et d'analyse d'images pour détecter automatiquement les signes de maladies sur les branches des pommiers. Ce système offre aux agriculteurs la possibilité de prendre des mesures préventives et de traitement plus rapides, réduisant ainsi les pertes de récolte et améliorant la durabilité de l'agriculture.

Les résultats obtenus par notre système ont démontré son efficacité dans la détection des maladies sur les branches des pommiers, ce qui confirme son potentiel à améliorer la santé des arbres fruitiers et à optimiser la production agricole. La combinaison de l'apprentissage automatique, du traitement d'image et de l'analyse des données a permis de développer un système capable d'identifier avec précision les anomalies et les signes de maladies. Les agriculteurs peuvent ainsi être alertés rapidement, leur permettant de prendre des mesures adéquates pour prévenir la propagation des maladies et réduire les pertes de récolte.

Cependant, il convient de souligner que notre système présente des opportunités d'amélioration et de recherche future. Une perspective intéressante consisterait à explorer de nouveaux ensembles de données pour la détection des maladies des pommiers. L'utilisation de nouveaux ensembles de données permettrait de tester et

de valider davantage notre système, en le confrontant à une plus grande variété de scénarios réels. Cela permettrait également d'améliorer sa robustesse et sa capacité à détecter un éventail plus large de maladies spécifiques aux pommiers.

De plus, une comparaison approfondie avec des algorithmes de deep learning tels que les réseaux de neurones convolutionnels (CNN) pourrait également être envisagée. Les CNN sont largement utilisés dans le domaine de la vision par ordinateur et ont montré de bonnes performances dans la classification et la détection d'objets. Comparer notre algorithme avec des approches basées sur les CNN nous permettrait d'évaluer la performance et la précision de notre système par rapport à d'autres méthodes couramment utilisées, ouvrant ainsi la voie à de nouvelles améliorations et à des recommandations pour l'application pratique de la détection des maladies des pommiers.

En conclusion, l'utilisation de l'IA pour la détection des maladies des pommiers présente un fort potentiel pour améliorer la santé des arbres fruitiers et optimiser la production agricole. Notre système intelligent basé sur l'IA constitue une avancée significative dans ce domaine, offrant une solution automatisée et précise pour la détection précoce des maladies sur les branches des pommiers. Cependant, des travaux futurs, notamment l'exploration de nouveaux ensembles de données et la comparaison avec des algorithmes de deep learning, sont nécessaires pour continuer à améliorer les performances et la précision de la détection des maladies sur les branches des pommiers. Ces développements contribueront à renforcer les pratiques agricoles durables et à soutenir les agriculteurs dans la préservation de leurs cultures, tout en offrant de nouvelles perspectives de recherche pour une utilisation encore plus efficace de l'intelligence artificielle dans l'agriculture moderne.

BIBLIOGRAPHIE

- [Alq+22] Sara ALQETHAMI et al. « Disease Detection in Apple Leaves Using Image Processing Techniques ». In : *Engineering, Technology & Applied Science Research* 12.2 (2022), p. 8335-8341.
- [Azi+19] Elaouaber Zineb AZIZA et al. « Decision tree CART algorithm for diabetic retinopathy classification ». In : *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)*. IEEE. 2019, p. 1-5.
- [BBW01] Richard BLUNDELL, Stephen BOND et Frank WINDMEIJER. *Estimation in dynamic panel data models : improving on the performance of the standard GMM estimator*. T. 15. Emerald Group Publishing Limited, 2001.
- [Bis19] Ekaba BISONG. *Building Machine Learning and Deep Learning Models on Google Cloud Platform : A Comprehensive Guide for Beginners*. https://www.researchgate.net/publication/336113323_Building_Machine_Learning_and_Deep_Learning_Models_on_Google_Cloud_Platform_A_Comprehensive_Guide_for_Beginners. En ligne; consulté le 20 mai 2023. 2019.
- [BK08] Gary BRADSKI et Adrian KAEHLER. *Learning OpenCV : Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.
- [BKA19] Saraansh BARANWAL, Siddhant KHANDELWAL et Anuja ARORA. « Deep learning convolutional neural network for apple leaves disease detection ».

- In : *Proceedings of international conference on sustainable computing in science, technology and management (SUSCOM), Amity University Rajasthan, Jaipur-India*. 2019.
- [BKG20] Dor BANK, Noam KOENIGSTEIN et Raja GIRYES. « Autoencoders ». In : *arXiv preprint arXiv :2003.05991* (2020).
- [CB21] Mayank Arya CHANDRA et SS BEDI. « Survey on SVM and their application in image classification ». In : *International Journal of Information Technology* 13 (2021), p. 1-11.
- [CR93] Leon O CHUA et Tamas ROSKA. « The CNN paradigm ». In : *IEEE Transactions on Circuits and Systems I : Fundamental Theory and Applications* 40.3 (1993), p. 147-156.
- [Dan92] Pierre-Auguste DANGEARD. *Les maladies du pommier et du poirier...* Librairie J.-B. Bailliere, 1892.
- [De 13] Barry DE VILLE. « Decision trees ». In : *Wiley Interdisciplinary Reviews : Computational Statistics* 5.6 (2013), p. 448-455.
- [DG+20] KERBAL DJAMILA, CHAIB BACHA GHANIA et al. « Étude de l'état phytopathologique des arbres fruitiers pépins (de poirier et de pommier) dans la wilaya d'Ain Defla ». In : (2020).
- [DLS95] FM DEWEY, R LI et T SWINBURNE. « A monoclonal antibody immunoassay for the detection of *Nectria galligena* in apple fruit and woody tissues 1 ». In : *Eppo bulletin* 25.1-2 (1995), p. 65-73.
- [Foo21] FOOD AND AGRICULTURE ORGANIZATION. *FAO - Food and Agriculture Organization of the United Nations*. Food et Agriculture Organization. 2021. URL : <https://www.fao.org/home/fr> (visité le 20/04/2023).
- [HM14] Meriem HAFFAF et Hanane MERZOUGUI. « Etude technico-culturelle du pommier *Malus pumila* Mill dans la région de M'sila ». Thèse de doct. Université Mohamed BOUDIAF de M'Sila, 2014.

- [Hss+14] Badr HSSINA et al. « A comparative study of decision tree ID3 and C4. 5 ». In : *International Journal of Advanced Computer Science and Applications* 4.2 (2014), p. 13-19.
- [Jam+09] Laurent JAMAR et al. « Protection contre la tavelure du pommier ciblée sur les infections primaires en production biologique ». In : *Actes des Journées techniques nationales fruits et légumes biologiques, 8-9 décembre 2009, ITAB-GRAB, Paris* (2009), p. 49-54.
- [JDF09] Chen JIN, Luo DE-LIN et Mu FEN-XIANG. « An improved ID3 decision tree algorithm ». In : *2009 4th international conference on computer science & Education*. IEEE. 2009, p. 127-130.
- [Joh17] Aayushi JOHARI. *Python Matplotlib Guide - Learn Matplotlib Library with Examples*. <https://medium.com/edureka/python-matplotlib-tutorial-15d148a7bfee>. En ligne ; consulté le 20 mai 2023. 2017.
- [Jup23] JUPYTER. *Jupyter*. <https://jupyter.org/>. En ligne ; consulté le 20 mai 2023. 2023.
- [Kha+22] Salman KHAN et al. « Transformers in vision : A survey ». In : *ACM computing surveys (CSUR)* 54.10s (2022), p. 1-41.
- [KQB21] Asif Iqbal KHAN, SMK QUADRI et Saba BANDAY. « Deep learning for apple diseases : classification and identification ». In : *International Journal of Computational Intelligence Studies* 10.1 (2021), p. 1-12.
- [Lec+01] Pascal LECOMPTE et al. « Pommier d'ornement et feu bactérien : de nombreux génotypes résistants ». In : *PHM Revue Horticole* 422 (2001), p. 58-60.
- [Lin+22] Tianyang LIN et al. « A survey of transformers ». In : *AI Open* (2022).
- [LR20] Xin LI et Laxmisha RAI. « Apple leaf disease identification and classification using resnet models ». In : *2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT)*. IEEE. 2020, p. 738-742.

- [MAL18] Purnawarman MUSA, Farid AL RAFI et Missa LAMSANI. « A Review : Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition ». In : *2018 third international conference on informatics and computing (ICIC)*. IEEE. 2018, p. 1-6.
- [Num23] NUMPY. *NumPy : the most used Python library in Data Science*. <https://datascientest.com/en/numpy-the-python-library-in-data-science>. En ligne ; consulté le 20 mai 2023. 2023.
- [Pan+19] Zhaoqing PAN et al. « Recent progress on generative adversarial networks (GANs) : A survey ». In : *IEEE access* 7 (2019), p. 36322-36333.
- [Per12] Sylvain PERICHON. « Les prés vergers de pommiers à cidre et les agroforêts de cacaoyers. Regard croisé sur l'évolution de l'arboriculture en Bretagne et en Amérique latine ». In : *Projets de paysage. Revue scientifique sur la conception et l'aménagement de l'espace* 7 (2012).
- [Per77] Robert PERRIN. « La lutte chimique contre le chancre du hêtre ». In : *Revue forestière française* 29.1 (1977), p. 27-32.
- [Pin+20] Walter Hugo Lopez PINAYA et al. « Autoencoders ». In : *Machine learning*. Elsevier, 2020, p. 193-208.
- [S1] S1. *Production mondiale de pommes par pays*. <https://www.atlasbig.com/fr-fr/pays-par-production-de-pommes>. Accessed : 2023-06-12.
- [S2] S2. *Khenchela : Une production de 1.46 million de quintaux de pommes durant la saison agricole 2018/2019* = <https://www.algerie-eco.com/2020/05/05/khenchela-une-production-de-1-46-million-q-de-pommes-durant-la-saison-agricole-2018-2019/>, note = Accessed : 2023-06-12.
- [SA22] Robin SONNARD et Flore ARALDI. « Protocole et suivi phytosanitaire du BioDiVerger (2017-2022) ». In : (2022).

- [SG99] Chris STAUFFER et W Eric L GRIMSON. « Adaptive background mixture models for real-time tracking ». In : *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. T. 2. IEEE. 1999.
- [SJ19] Kamilya SMAGULOVA et Alex Pappachen JAMES. « A survey on LSTM memristive neural network architectures and applications ». In : *The European Physical Journal Special Topics* 228.10 (2019), p. 2313-2324.
- [SS18] Pramila P SHINDE et Seema SHAH. « A review of machine learning and deep learning applications ». In : *2018 Fourth international conference on computing communication control and automation (ICCUBEA)*. IEEE. 2018, p. 1-6.
- [The] @THEPSF. *Python*. <https://www.python.org/doc/essays/blurb/>. En ligne; consulté le 20 mai 2023.
- [USS01] RS UTKHEDE, PL SHOLBERG et MJ SMIRLE. « Effects of chemical and biological treatments on growth and yield of apple trees planted in *Phytophthora cactorum* infected soil ». In : *Canadian Journal of Plant Pathology* 23.2 (2001), p. 163-167.
- [VM02] SVM VISHWANATHAN et M Narasimha MURTY. « SSVM : a simple SVM algorithm ». In : *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*. T. 3. IEEE. 2002, p. 2393-2398.
- [Yin+17] Wenpeng YIN et al. « Comparative study of CNN and RNN for natural language processing ». In : *arXiv preprint arXiv :1702.01923* (2017).
- [YY+20] Divakar YADAV, Arun Kumar YADAV et al. « A Novel Convolutional Neural Network Based Model for Recognition and Classification of Apple Leaf Diseases. » In : *Traitement du Signal* 37.6 (2020).
- [ZA14] Kawthar Yasmine ZERGAT et Abderrahmane AMROUCHE. « New scheme based on GMM-PCA-SVM modelling for automatic speaker recognition ». In : *International Journal of Speech Technology* 17 (2014), p. 373-381.