HAL
archives-ouvertes.fr

Analyse d'images pour l'identification multi-organes d'espèces végétales

Sarah Bertrand

▶ **To cite this version:**

Sarah Bertrand. Analyse d'images pour l'identification multi-organes d'espèces végétales. Environnements Informatiques pour l'Apprentissage Humain. Université de Lyon, 2018. Français. <NNT : 2018LYSE2127>. <tel-01951469>

HAL Id: tel-01951469
https://hal.archives-ouvertes.fr/tel-01951469

Submitted on 11 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ LUMIÈRE LYON 2
École Doctorale Informatique et Mathématiques de Lyon

THÈSE

pour obtenir le grade de
DOCTEUR EN INFORMATIQUE

présentée et soutenue par

Sarah BERTRAND

le 10 décembre 2018

Analyse d'images pour l'identification multi-organes d'espèces végétales

---

Devant le jury composé de :

Rapporteur : Pascal DESBARATS, Professeur, Université de Bordeaux

Rapportrice : Anne VERROUST-BLONDET, Chargée de Recherches, INRIA

Rapporteur : Christophe DUCOTTET, Professeur, Université Jean Monnet

Examinatrice : Jasmine BURGUET, Chargée de recherche, INRA

Examinateur : David ROUSSEAU, Professeur, Université d'Angers

Directrice de thèse : Laure TOUGNE, Professeure, Université Lyon 2

Co-directeur de thèse : Guillaume CERUTTI, Ingénieur de Recherches, ENS de Lyon

Remerciements

Mes remerciements vont bien évidemment en premier lieu à ma directrice de thèse Laure Tougne, professeure à l'Université Lumière Lyon 2, pour nos trois années et plus à travailler ensemble sur ce lourd projet qu'est une thèse. Elle a su répondre à mes questions avec patience et diriger mes travaux dans la bonne direction. La charge de travail n'a pas été équitablement répartie tout au long de ces années et je la remercie d'avoir su y répondre. Je remercie aussi mon co-directeur de thèse Guillaume Cerutti, ingénieur de recherche à l'ENS de Lyon, pour son expertise scientifique et ses phrases à rallonge.

Je tiens aussi à remercier les membres de mon jury d'avoir accepté de considérer les apports scientifiques présentés dans ce manuscrit. Je tiens à citer notamment Pascal Desbarats, professeur à l'Université de Bordeaux, Anne Verroust-Blondet, chargée de recherche à l'INRIA, et Christophe Ducottet, professeur à l'Université de Saint-Étienne, pour avoir accepté de rapporter ma thèse. Je remercie aussi Jasmine Burguet, chargée de recherche à l'INRA-AgroParisTech, et David Rousseau, professeur rattaché au laboratoire LARIS à Angers, de prendre part à mon jury en tant qu'examinateurs.

J'aimerais ensuite remercier tous mes collègues actuels et anciens du LIRIS Lyon 2, pour la bonne ambiance au travail, les discussions de jardinage, chats, lapin et poules, les profusions de sucreries et raclette-party bonnes pour l'esprit mais mauvaises pour le corps, et les partages d'astuces pour franchir les blocus. Un merci particulier à Dominique Maniez, pour avoir chassé les fautes d'orthographes, et Catherine Villeseche, pour sa disponibilité et réactivité pour les démarches administratives et les voyages en conférences.

Pour finir, je tiens à remercier ma famille, et plus précisément ma mère, qui a fait les trajets chaque week-end pour m'emmener à la gare afin que je puisse faire des études supérieures. Enfin, je remercie Jean-Rémi, pour avoir rechargé mes batteries quand j'en avais besoin.

Résumé

Cette thèse s'inscrit dans le cadre de l'ANR ReVeRIES dont l'objectif est d'utiliser les technologies mobiles pour aider l'homme à mieux connaître son environnement et notamment les végétaux qui l'entourent. Plus précisément, le projet ReVeRIES s'appuie sur une application mobile, nommée *Folia* développée dans le cadre du projet ANR ReVeS, capable de reconnaître les espèces d'arbres et arbustes à partir de photos de leurs feuilles. Ce prototype se différencie des autres outils car il est capable de simuler le comportement du botaniste. Dans le contexte du projet ReVeRIES, nous nous proposons d'aller beaucoup plus loin en développant de nouveaux aspects : la reconnaissance multimodale d'espèces, l'apprentissage par le jeu et les sciences citoyennes. L'objet de cette thèse porte sur le premier de ces trois aspects, à savoir l'analyse d'images d'organes de végétaux en vue de l'identification. Plus précisément, nous considérons les principaux arbres et arbustes, endémiques ou exotiques, que l'on trouve en France métropolitaine. L'objectif de cette thèse est d'étendre l'algorithme de reconnaissance en prenant en compte d'autres organes que la feuille. Cette multi-modalité est en effet essentielle si nous souhaitons que l'utilisateur apprenne et s'entraîne aux différentes méthodes de reconnaissance, pour lesquelles les botanistes utilisent la variété des organes (*i.e.* les feuilles, les fleurs, les fruits et les écorces). La méthode utilisée par *Folia* pour la reconnaissance des feuilles étant dédiée, car simulant le botaniste, ne peut s'appliquer directement aux autres organes. Ainsi, de nouveaux verrous se posent, tant au niveau du traitement des images qu'au niveau de la fusion de données. Une première partie de la thèse a été consacrée à la mise en place de méthodes de traitement d'images pour l'identification des espèces végétales. C'est l'identification des espèces d'arbres à partir d'images d'écorces qui a été étudiée en premier. Les descripteurs développés prennent en compte la structure de l'écorce en s'inspirant des critères utilisés par les botanistes. Les fruits et les fleurs ont nécessité une étape de segmentation avant leur description. Une nouvelle méthode de segmentation réalisable sur smartphone a été développée pour fonctionner sur la grande variabilité des fleurs et des fruits. Enfin, des descripteurs ont été extraits sur les fruits et les fleurs après l'étape de segmentation. Nous avons décidé de ne pas faire de séparation entre les fleurs et les fruits car nous avons montré qu'un utilisateur novice en botanique ne sait pas toujours faire la différence entre ces deux organes sur des arbres dits « d'ornement » (non fruitiers). Pour les fruits et les fleurs, la prédiction n'est pas seulement faite sur les espèces mais

aussi sur les genres et les familles, groupes botaniques traduisant d'une similarité entre ces organes. Une deuxième partie de la thèse traite de la combinaison des descripteurs des différents organes que sont les feuilles, les écorces, les fruits et les fleurs. En plus des méthodes de combinaison basiques, nous proposons de prendre en compte la confusion entre les espèces, ainsi que les prédictions d'appartenance aux taxons botaniques supérieurs à l'espèce. Enfin, un chapitre d'ouverture est consacré au traitement de ces images par des réseaux de neurones à convolutions. En effet, le Deep-Learning est de plus en plus utilisé en traitement d'images, notamment appliqué aux organes végétaux. Nous proposons dans ce contexte de visualiser les filtres de convolution extrayant de l'information, afin de faire le lien entre les informations extraites par ces réseaux et les éléments botaniques.

Mots clefs : Reconnaissance de plantes, segmentation, extraction de descripteurs, classification, fusion de données, réseaux de neurones.

Abstract

This thesis is part of the ANR ReVeRIES, which aims to use mobile technologies to help people better understand their environment and in particular the plants that surround them. More precisely, the ReVeRIES project is based on a mobile application called *Folia* developed as part of the ANR ReVeS project and capable of recognising tree and shrub species based on photos of their leaves. This prototype differs from other tools in that it is able to simulate the behaviour of the botanist. In the context of the ReVeRIES project, we propose to go much further by developing new aspects : multimodal species recognition, learning through play and citizen science. The purpose of this thesis is to focus on the first of these three aspects, namely the analysis of images of plant organs for identification.

More precisely, we consider the main trees and shrubs, endemic or exotic, found in metropolitan France. The objective of this thesis is to extend the recognition algorithm by taking into account other organs in addition to the leaf. This multi-modality is indeed essential if we want the user to learn and practice the different methods of recognition for which botanists use the variety of organs (*i.e.* leaves, flowers, fruits and bark). The method used by *Folia* for leaf recognition being dedicated, because simulating the work of a botanist on the leaf, cannot be applied directly to other organs. Thus, new challenges are emerging, both in terms of image processing and data fusion.

The first part of the thesis was devoted to the implementation of image processing methods for the identification of plant species. The identification of tree species from bark images was the first to be studied. The descriptors developed take into account the structure of the bark inspired from the criteria used by botanists. Fruits and flowers required a segmentation step before their description. A new segmentation method that can be used on smartphones has been developed to work in spite of the high variability of flowers and fruits. Finally, descriptors were extracted on fruits and flowers after the segmentation step. We decided not to separate flowers and fruits because we showed that a user new to botany does not always know the difference between these two organs on so-called "ornamental" trees (not fruit trees). For fruits and flowers, prediction is not only made on their species but also on their genus and family, botanical groups reflecting a similarity between these organs. The second part of the thesis deals with the combination of descriptors of the different organs : leaves, bark, fruits and flowers. In addition to

basic combination methods, we propose to consider the confusion between species, as well as predictions of affiliations in botanical taxa higher than the species. Finally, an opening chapter is devoted to the processing of these images by convolutional neural networks. Indeed, Deep Learning is increasingly used in image processing, particularly for plant organs. In this context, we propose to visualize the learned convolution filters extracting information, in order to make the link between the information extracted by these networks and botanical elements.

KEYWORDS : Plant recognition, segmentation, descriptor extraction, classification, data fusion, neural networks.

Table des matières

Introduction générale **1**

1 Éléments botaniques **9**

1.1 Classification botanique des végétaux 11

1.1.1 Taxinomie 11

1.1.2 Nomenclature des espèces 12

1.2 Identification botanique des Angiospermes 15

1.2.1 Clefs de détermination botaniques 15

1.2.2 Identification via les fleurs 17

1.2.3 Identification via les fruits 20

1.2.4 Identification via les feuilles 22

1.2.5 Identification via les écorces 26

1.3 Reconnaissance de végétaux par des novices 28

1.3.1 Questionnaire de différenciation des fruits et des fleurs 29

1.4 Utilisation de clefs de détermination pour novices 32

1.4.1 Clefs "papier" 33

1.4.2 Clefs numériques 33

2 État de l'art **37**

2.1 Systèmes automatiques de reconnaissance de végétaux 39

2.1.1 Applications basées sur l'indexation d'images par le contenu 39

2.1.2 Applications basées sur le traitement d'images 40

2.2 Détection et segmentation des différents organes de l'arbre 43

2.2.1 Segmentation de l'utilisateur 43

2.2.2 Segmentation automatique 47

2.3 Extraction des descripteurs 52

2.3.1 Descripteurs de forme 52

2.3.2 Descripteurs de couleur et de texture 53

2.3.3 Descripteurs locaux 56

2.4 Classification 56

2.5 Les réseaux de neurones à convolutions (CNN) 58

2.6 Cadre de travail . . . 59
2.6.1 Segmentation d'organes végétaux en milieu naturel complexe . . . 60
2.6.2 Description et classification/fusion des informations basées botanique des différents organes . . . 61
2.6.3 Bases de données . . . 62

3 Reconnaissance d'écorces 65
3.1 Cadre de travail . . . 67
3.1.1 Évolution biologique de l'écorce . . . 67
3.1.2 Base de données utilisée . . . 69
3.2 Extraction des caractéristiques . . . 70
3.2.1 Orientation principale de l'écorce . . . 70
3.2.2 Description de la texture . . . 72
3.2.3 Caractérisation de la couleur de l'écorce . . . 75
3.3 Classification . . . 77
3.3.1 SVM un contre tous . . . 77
3.3.2 Classifeur Perceptron Multi-couches (MLP) . . . 79
3.4 Influence des différents descripteurs . . . 80
3.4.1 Influence du vecteur Gabor . . . 80
3.4.2 Influence des périodes du quadrillage pour l'extraction du "mot vertical" et du "mot horizontal" . . . 81
3.4.3 Réduction de l'histogramme . . . 83
3.5 Comparaison . . . 87
3.5.1 Comparaison sur la base de données *NewBarkTex* . . . 87
3.5.2 Comparaison avec un descripteur de texture . . . 89
3.5.3 Comparaison avec le Deep Learning . . . 90
3.6 Conclusion . . . 92

4 Segmentation des fruits et fleurs 93
4.1 Cadre de travail . . . 95
4.1.1 Création de la base de données . . . 97
4.2 Carte de MLP . . . 100
4.2.1 Carte de couleurs . . . 103
4.2.2 Performance de la carte de MLP . . . 106
4.2.3 Influence de la trace utilisateur . . . 108
4.3 Carte de profondeurs . . . 110
4.3.1 Influence de la trace utilisateur . . . 117
4.3.2 Construction de la carte avec un même appareil photo . . . 119
4.4 Contour actif . . . 123
4.4.1 Comparaison sur une base de données de feuille . . . 125

4.4.2 Contribution de la carte de profondeurs . . . 126
4.4.3 Gestion des poids des deux cartes . . . 126
4.5 Conclusion . . . 128

5 Classification des fruits et fleurs 131
5.1 Cadre de travail . . . 133
5.1.1 Base de données . . . 133
5.1.2 Préparation des données . . . 134
5.1.3 Fusion des inflorescences et des fruits . . . 135
5.2 Descripteurs . . . 135
5.2.1 Descripteur de couleur . . . 135
5.2.2 Descripteur de forme . . . 137
5.2.3 Descripteur de texture . . . 139
5.3 Classifieur . . . 141
5.3.1 Augmentation des données . . . 141
5.3.2 Classifieur SVM . . . 144
5.3.3 Classifieur MLP . . . 145
5.4 Comparaison avec des méthodes de Deep Learning . . . 153
5.4.1 Images non segmentées . . . 154
5.4.2 Images segmentées . . . 156
5.5 Conclusion . . . 157

6 Stratégies de fusion des données pour la classification multi-organes d'espèces végétales 159
6.1 Identification des feuilles dans l'application *Folia* . . . 161
6.1.1 Descripteurs de feuilles . . . 161
6.1.2 Classifieur de feuilles . . . 162
6.2 Base de données . . . 162
6.3 Méthodes de combinaisons basiques . . . 163
6.3.1 Concaténation des vecteurs . . . 163
6.3.2 Combinaison *a posteriori* . . . 164
6.4 Prise en compte de la confusion entre les espèces . . . 166
6.4.1 Matrices de confusion issues de la classification des feuilles et la classification des écorces . . . 166
6.4.2 Matrices de confusion pour la concaténation et de la combinaison *a posteriori* des vecteurs de feuilles et des écorces . . . 169
6.4.3 Prise en compte des matrices de confusion . . . 171
6.5 Ajout des données des fruits et des fleurs . . . 177
6.5.1 Raffinement de la prédiction de l'espèce . . . 178
6.5.2 Combinaison *a posteriori* des fruits-fleurs avec les feuilles et les écorces 178

6.5.3 Prise en compte du genre et non de l'espèce 181
6.6 Conclusion 183

7 Interprétation sémantique des informations extraites par un réseau convolutionnel 185
7.1 Cadre de travail 187
7.1.1 Réseau de neurones à convolutions utilisé 187
7.1.2 Données utilisées 189
7.1.3 Performance du CNN 190
7.2 Visualisation des couches 190
7.2.1 Principes et fonctionnement 190
7.2.2 Visualisation préliminaire 192
7.3 Détection automatique des filtres actifs 196
7.3.1 Méthodes de traitement d'images basiques en vue de l'automatisation 197
7.3.2 Méthode basée sur le seuillage d'Otsu 199
7.4 Identification des filtres révélateurs d'une caractéristique botanique 202
7.5 Conclusion 205

Conclusion générale 207

Bibliographie 215

Introduction générale

Contexte

Le nombre d'espèces appartenant au règne végétal est difficile à déterminer. Il existerait plus de 400 000 espèces, et en moyenne plus de 2 000 espèces sont découvertes par an [Wil17]. Les plantes ont été étudiées depuis les premiers temps pour leurs propriétés médicinales, nutritives, aromatiques, horticoles, ou matérielles. Au moins 28 187 espèces végétales sont actuellement déclarées comme étant à usage médicinal [Wil17]. En France, la flore est estimée à plus de 5 000 espèces de plantes. Cependant, avec la croissance de l'urbanisation et des technologies, les êtres humains ne sont aujourd'hui plus en phase avec les végétaux, et ont oublié les noms des plantes et leurs usages.

Néanmoins, une prise de conscience est en train de naître concernant l'environnement et ses ressources naturelles. L'identification des espèces de plantes (savoir les nommer) semble indispensable pour comprendre l'environnement végétal. Cependant, cette tâche peut être ardue, les connaissances liées à l'identification des espèces étant souvent connues uniquement des botanistes, des scientifiques, ou de quelques citoyens intéressés. Elles sont généralement incomplètes pour des environnements qui ont une grande diversité de plantes. De ce fait, l'identification des espèces végétales par des néophytes est quasiment impossible, et parfois même difficile pour des professionnels tels que des fermiers ou des gérants de forêts, ce qui est dénommé fossé taxonomique.

Ainsi, fournir à une personne novice, ne possédant aucun bagage botanique, un outil permettant de connaître les plantes qui l'entourent, et lui apprendre à les reconnaître, serait une grande avancée. En plus du fait de pouvoir nommer les plantes, un tel outil pourrait servir dans le cas de la détection d'espèces invasives (6075 déclarées en 2017 [Wil17]).

D'autre part, de plus en plus de gens possèdent un téléphone portable de type smartphone. Ces smartphones incluent une multitude de fonctionnalités telles qu'un appareil photo, une caméra, un GPS, un accéléromètre, une connexion Internet, *etc.* L'utilisateur a également la possibilité d'installer de nouvelles applications de tout type qu'il pourrait utiliser dans la vie quotidienne. Ces smartphones sont donc des supports idéaux pour les applications liées à l'analyse d'images, permettant des interactions avec l'utilisateur, mais avec certaines contraintes. Une application sur smartphone requiert un espace de stockage et une puissance de calcul limités, et en termes de traitement d'images, l'environnement de prise de vue est non contrôlé (luminosité, position, *etc*).

Nous cherchons donc à développer un outil automatique, ou semi-automatique, sur smartphone pour enseigner ou aider un utilisateur à reconnaître des espèces végétales.

Projet ReVeRIES (ANR15-CE38-004-01)

Cette thèse s'inscrit dans le cadre du projet ANR ReVeRIES (Reconnaissance de Végétaux Récréative, Interactive et Éducative sur Smartphone), dont le but premier est de rapprocher l'homme et la nature en rendant plus abordable la botanique. Ce projet s'appuie sur une application mobile, dénommée Folia [1] et développée dans le cadre du projet ANR ReVeS (ANR-10-CORD-005). Elle est capable de reconnaître les espèces d'arbres et d'arbustes de plus de 1m20 originaires de France en analysant des photos de leurs feuilles. Ce prototype simule le comportement d'un botaniste en essayant de déterminer l'espèce de la plante en extrayant des caractéristiques botaniques, ce qui la différencie de tous les autres outils disponibles sur le marché (Figure 1). Les feuilles ont été étudiées pour la reconnaissance des espèces d'arbres car elles peuvent être accessibles quasiment tout au long de l'année. De plus, ce sont ces organes qui sont principalement regardés lorsque des notions de botaniques sont enseignées pour la reconnaissance des espèces d'arbres.

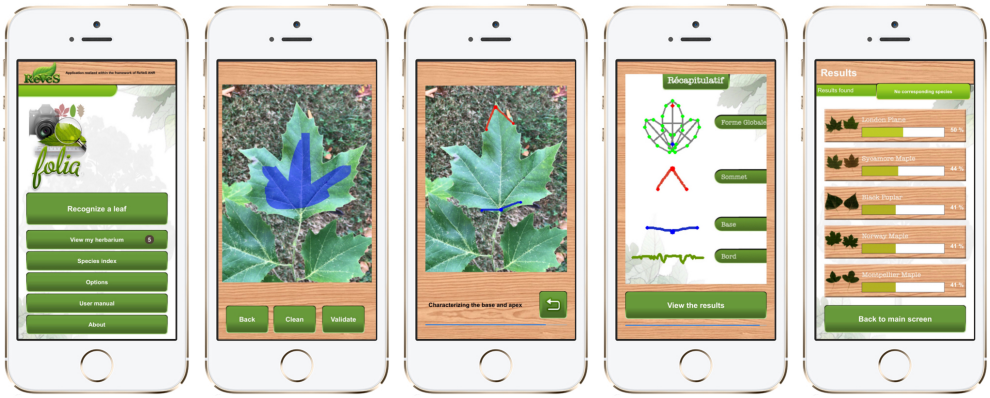


Figure 1 – Illustrations de différentes interfaces de l'application Folia de reconnaissance d'espèces d'arbres via une photo de feuille sur smartphone.

Dans le cadre du projet ReVeRIES, l'apprentissage mobile basé sur le jeu, la reconnaissance multi-organes et la science citoyenne sont enrichis. Ses principaux objectifs sont donc :

— L'inclusion de mécanismes de jeux appropriés et adaptatifs pour susciter l'apprentissage, la motivation et l'engagement. Ces mécanismes aideront les utilisateurs à se familiariser avec les caractéristiques des plantes et surtout à apprendre les méthodes utilisées par des botanistes experts pour reconnaître les familles, genres et espèces de plantes. Les utilisateurs pourront améliorer leurs compétences en comparant leurs résultats à ceux de l'algorithme de reconnaissance.

— La création d'algorithmes de reconnaissance semi-automatique d'espèces d'arbres et d'arbustes à partir d'images naturelles de leurs organes (feuilles, écorce, fruits, fleurs) simulant l'action du botaniste. L'utilisation de plusieurs organes devrait

1. Application gratuite pour téléphone Apple et Android http ://liris.univ-lyon2.fr/reves/folia/public/

améliorer la précision de l'algorithme par rapport aux feuilles seules. Cette thèse se situe dans cet objectif.

— Le développement d'inventaires collaboratifs de plantes (science citoyenne). Les photos et informations géolocalisées, prises avec l'application et validées par des experts, pourraient être transférées à des réseaux spécialisés, tels que Tela Botanica [2], intégrés dans le système d'information géographique OpenStreetMap [3] et mobilisés par les institutions locales pour soutenir des actions et des projets impliquant les citoyens.

Globalement, le projet soulève de nombreux défis scientifiques dans les domaines des TEL (Technology Enhanced Learning), de l'analyse d'images, de la fusion de données, et aussi dans le domaine des inventaires environnementaux collaboratifs. Les impacts possibles sont nombreux : enseignement de la botanique à différents niveaux et auprès de publics d'apprentissage variés, intelligence collective, science citoyenne, conservation de la nature et jeux collaboratifs environnementaux. En plus des citoyens intéressés par la nature, ce système pourrait être très utile pour les enseignants et leurs élèves, les botanistes et les parcs naturels.

Le projet ReVeRIES regroupe cinq laboratoires de recherche partageant leurs expertises pour le mener à bien :

— Le LIUM (Laboratoire d'Informatique de l'Université du Maine) à Laval (EA 4023), apportant des compétences dans l'apprentissage par le jeu (serious gaming) et la modélisation des apprenants ;

— Le LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) à Lyon (UMR 5205), dont l'expertise se situe au niveau du traitement d'image et de l'extraction de descripteurs ;

— Le LISTIC (Laboratoire d'Informatique, Systèmes, Traitement d'Information et de la Connaissance) à Annecy (EA 3703) spécialisé en traitement et en fusion d'informations.

— EVS (Environnement, Ville, Société) à Saint-Étienne (UMR 5600), laboratoire de géographie apportant notamment une compétence en géomatique et systèmes d'information géographiques ;

— L'IRHS (Institut de Recherche en Horticulture et Semences) à Angers (UMR 1345) apportant une expertise en taxonomie et nomenclature botanique.

Cette thèse au sein du laboratoire LIRIS s'inscrit donc dans le domaine de l'analyse d'image pour l'identification multi-organe d'espèce d'arbre et arbustes. La problématique de la reconnaissance des espèces d'arbres via les feuilles ayant déjà été traitée dans le cadre du projet ReVeS, nous nous intéressons dans cette thèse à la reconnaissance d'espèces grâce

2. http ://www.tela-botanica.org, réseau francophone de référence de la communauté botanique.

3. https ://openstreetmap.fr, projet international ayant pour but de créer une carte libre du monde de manière collaborative.

aux autres organes de l'arbre que sont le tronc (et plus particulièrement l'écorce), les fruits et les fleurs.

Plan du manuscrit

L'organisation de ce mémoire est la suivante.

Le chapitre 1 présentera les méthodes de classification et de reconnaissance des végétaux utilisées par les botanistes ; méthodes sur lesquelles sont basés certains des algorithmes de reconnaissance. Il traitera aussi de la position du novice par rapport au vocabulaire botanique et des applications déjà existantes qui lui sont proposées pour répondre à ses demandes.

Le chapitre 2 fera l'état de l'art des différentes applications pour smartphone de reconnaissance de végétaux et de leur fonctionnement, et détaillera certaines publications scientifiques pour le traitement d'images associées à la reconnaissance de végétaux.

Le chapitre 3 traitera de la caractérisation des écorces et de leur classification.

Le chapitre 4 décrira quant à lui la méthode de segmentation conçue pour l'extraction des fruits et des fleurs en milieu complexe (sur fond naturel) à partir d'informations de texture, couleur et de profondeur.

Le chapitre 5 fera suite au chapitre antérieur en exposant le procédé de reconnaissance de fruits et de fleurs.

Le chapitre 6 détaillera le procédé de fusion de données des vecteurs caractéristiques des feuilles utilisés dans l'application Folia et des vecteurs caractéristiques des écorces décrits dans le chapitre précédent.

Enfin, le chapitre 7 sera une extension du travail de thèse en abordant la reconnaissance de végétaux grâce aux réseaux de neurones à convolutions (CNN), et plus particulièrement la visualisation des caractéristiques extraites par ceux-ci en fonction des différentes couches.

Contributions apportées

Les apports scientifiques développés au cours de cette thèse et détaillés dans les différents chapitres de ce manuscrit ont fait l'objet de plusieurs communications dans des conférences et journaux à comité de relecture international.

Le travail sur la caractérisation des écorces, leur classification et leur apport par rapport à la reconnaissance d'espèces d'arbres à partir des feuilles a été présenté à Porto (Portugal) en février 2017 lors de la conférence VISAPP (International Conference on Computer Vision Theory and Applications) [BCT17a].

La méthode de segmentation des fruits et des fleurs en milieu naturel a été exposée lors de la conférence ICMV (International Conference on Machine Vision) à Vienne (Autriche)

en novembre 2017 [BCT17b].

Avec les informations extraites sur les feuilles et sur les écorces, un article de journal a été rédigé conjointement avec nos partenaires de projet du laboratoire LISTIC comparant différentes méthodes de fusion de données adaptées aux données de feuilles et d'écorces, dont une que nous avons proposée prenant en compte la confusion entre espèces. Cet article a été publié en juillet 2018 dans la revue Ecological Informatics [Ber+18].

Pour finir, l'étude de la visualisation des couches d'un réseau de neurones à convolutions (CNN) entraîné sur une base de données de feuilles, présenté dans le Chapitre 7, a été exposée lors de la conférence IAMPS (International Workshop on Image Analysis Methods for the Plant Sciences) à Nottingham (Royaume-Uni) en janvier 2018 [BCT18b].

Chapitre 1
Éléments botaniques

L'objectif de ce chapitre est de donner un peu d'outils et de vocabulaire botanique pour comprendre les différentes contraintes que pose l'identification des végétaux. Il y sera exposé les méthodes botaniques de classification et de dénomination des espèces. Ce chapitre montrera que chaque organe de la plante peut donner des indices pour permettre d'identifier une espèce, mais que tous ne se valent pas. Dans une seconde partie, des outils spécialisés pour les botanistes seront confrontés à leur utilisation par des novices, et des solutions seront énoncées.

1.1 Classification botanique des végétaux

Définissons dans un premier temps ce qu'est l'identification botanique. La plupart des novices savent reconnaître des plantes. Reconnaître ne requiert pas le besoin de nommer une plante, mais de juger si elle a déjà été vue. Nous avons déjà vu un arbre, de l'herbe. Or, l'identification demande quant à elle de nommer précisément une plante. Par exemple, l'arbre observé est un chêne pédonculé (nom scientifique *Quercus robur*), l'herbe est du vulpin des prés (nom scientifique *Alopecurus pratensis*) (Figure 1.1).

FIGURE 1.1 – Photo d'un chêne pédonculé et photo de plants de vulpins des prés (épis).

En botanique, l'identification d'une plante consiste à pouvoir trouver ou retrouver le nom de celle-ci, soit de mémoire, soit à l'aide d'outils botaniques. Nous allons voir que cette tâche peut être complexe pour un novice.

1.1.1 Taxinomie

Les plantes, comme tous les êtres vivants, sont réunies en groupes, appelés taxons, selon une hiérarchie. La systématique est la science qui organise ce classement et qui étudie les relations entre les taxons. La systématique comprend aussi la taxinomie qui, en biologie, a pour but de décrire les organismes vivants et de les regrouper en taxons afin de les reconnaître, les identifier, les nommer, et les classer avec différents caractères (morphologiques, génétiques) grâce à des outils que sont les clés de détermination. Cette connaissance de la classification des espèces et la manière dont elles sont ordonnées entre elles pourra nous aider dans l'établissement de descripteurs et de méthodes de classification pertinents.

1.1.2 Nomenclature des espèces

Commençons par définir la notion d'espèce. L'espèce, qui permet de distinguer les différents types d'organismes vivants, est un concept ambigu qui possède plusieurs définitions dans la littérature. La définition la plus utilisée est celle de Ernst Mayr :

> *Species are groups of actually or potentially interbreeding populations, which are reproductively isolated from other such groups*
>
> Ernst Mayr, 1942

Il a plus tard été ajouté à cette définition le fait que ce groupe produise une descendance viable et féconde. Ainsi, l'espèce est le plus grand ensemble de populations au sein de laquelle le flux génétique est possible dans des conditions naturelles, les individus d'une même espèce étant génétiquement isolés d'autres ensembles équivalents du point de vue reproductif.

Cependant, il existe dans le règne végétal de nombreuses espèces qui s'hybrident naturellement, au sein d'un même genre comme les bouleaux (genre *Betula*) ou les chênes (genre *Quercus*), sans que les taxonomistes les considèrent comme une seule espèce. Le critère reproductif est donc un peu difficile d'application dans l'ensemble des groupes, bien que généralement considéré comme décisif pour distinguer deux espèces.

Actuellement, les êtres vivants sont agencés dans une classification dite hiérarchique, et comprenant divers rangs (cf. Figure 1.2, où les rangs sont nommés). Par ailleurs, la pratique des systématiciens actuels vise à faire correspondre l'histoire des organismes (leur apparentement, leur phylogénie) avec la classification. La phylogénie des organismes est étudiée selon une approche cladistiques, conduisant à la production d'arbres phylogénétiques, et certains clades de ces arbres phylogénétiques sont reconnus comme des taxons de rang X ou Y dans la classification.

Chaque espèce porte un nom binomial, une combinaison de deux termes latins, dont le premier correspond au nom générique (le nom de genre) et le second à l'épithète spécifique. Par exemple dans *Quercus robur*, "*Quercus*" est le nom de genre, "*robur*" est l'épithète spécifique, la combinaison des deux, *i.e. Quercus robur*, étant le nom scientifique du chêne pédonculé". Le nom binomial permet de désigner de manière unique une espèce, contrairement au nom commun, dit nom vernaculaire, qui n'est pas forcément précis et qui peut représenter plusieurs espèces. Nous pouvons prendre l'exemple du nom vernaculaire "chêne", qui peut désigner le chêne vert (*Quercus ilex*), le chêne pédonculé (*Quercus robur*), le chêne-liège (*Quercus suber*), *etc.* Par ailleurs, une espèce peut avoir plusieurs noms vernaculaires en fonction de la région ou du patois local. Par exemple, *Quercus suber* est appelé chêne-liège, corcier, surier ou suve.

Le genre est un rang taxinomique qui traduit une proximité d'espèces dans leurs aspects (morphologiques, génétiques, *etc.*), bien que discernables entre elles. Par exemple

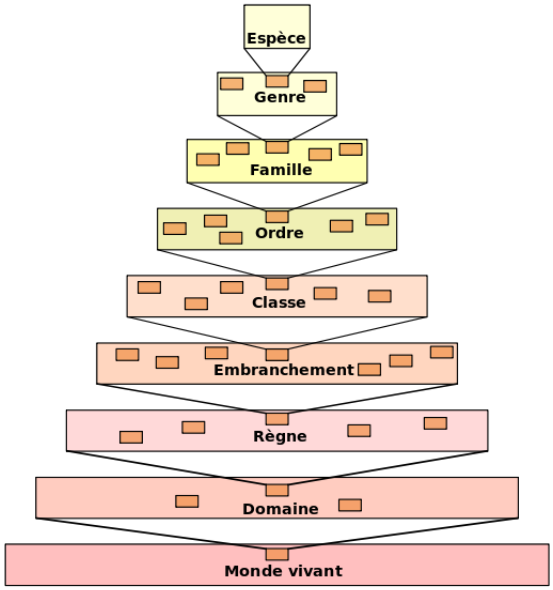


FIGURE 1.2 – Hiérarchie taxonomique du vivant dans la classification classique (CC BY-SA 3.0 via Wikimedia Commons)

au sein d'un même genre, les espèces produisent les mêmes types de fruits. En effet, le genre *Fagus* donne des faînes, le genre *Quercus* des glands, le genre *Castanea* des châtaignes (Figure 1.3). C'est aussi généralement au sein d'un même genre que vont se produire les hybridations alors appelées interspécifiques (par exemple le casseillier *Ribes nigrum* x *Ribes uva-crispa* est un hybride entre le cassissier et le groseillier à maquereau).

FIGURE 1.3 – Les espèces du genre *Quercus* donnent des glands (ici *Quercus cerris*, *Quercus coccifera*, *Quercus ilex*, *Quercus robur*).

Le taxon supérieur au genre est la famille. La terminaison latine des familles chez les plantes est *-aceae*, traduit en –acées en français. Ce taxon regroupe les genres qui sont les plus proches et qui rassemble des espèces qui visuellement commencent à être très différentes les unes des autres. Ainsi, les genres *Carpinus* (charmes) *Alnus* (aulnes), *Betula* (bouleaux) et *Corylus* (noisetiers) appartiennent à la famille *Betulaceae*. Ce sont classiquement les organes reproducteurs des plantes (fleurs) qui possèdent des propriétés morphologiques communes qui sont utilisés pour rassembler des genres dans une même famille (Figure 1.4). Les fruits issus de ces fleurs fécondées diffèrent cependant au sein d'une même famille (Figure 1.5).

FIGURE 1.4 – Fleurs de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae*

FIGURE 1.5 – Fruits de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae*

Il existe des cas d'hybridations intergénériques, par exemple le limequat, hybridation entre un kumquat (*Fortunella japonica*) et un limetier (*Citrus aurantiifolia*), mais toujours au sein d'une même famille (*Rutaceae* dans notre exemple).

Les taxons supérieurs correspondent à de grands clades divisant les organismes vivants. L'ordre est une notion peu claire qui regroupe plusieurs familles. Les plantes à fleurs ont été traditionnellement divisées en 2 classes en deux groupes : les *Liliopsida* (monocotylédones, dont les jeunes pousses ne possèdent qu'une seule feuille primordiale) et les *Magnoliopsida* (dicotylédones, dont les jeunes pousses ne possèdent que deux feuilles primordiales). L'embranchement caractérise le type de reproduction, par exemple l'embranchement des Angiospermes (les plantes à fleurs dont la graine est protégée) ou les Gymnospermes (qui ont une graine nue et regroupant l'ensemble des conifères). Dans notre Ère actuelle, les chiffres montrent la domination qu'exercent aujourd'hui les Angiospermes parmi les plantes (90 % de la diversité végétale en dehors des océans). Nous allons donc nous intéresser exclusivement à cette division dans cette thèse. Enfin, le règne est le niveau le plus haut de classification des êtres vivants. Les plantes font partie du règne végétal *Plantae*. Pour résumer, si nous prenons l'exemple du figuier, dont le nom binomial est *Ficus carica*, l'épithète spécifique est *carica*, le genre *Ficus*, la famille *Moraceae*, appartenant à l'ordre des *Urticales*, inclus dans la classe des *Magnolipsida*.

Il existe d'autres systèmes de classification, et les théories de l'évolution ont contrarié l'agencement de la classification classique. La classification phylogénétique, ou classifica-

tion évolutive, a pour objectif de rendre compte des degrés de parenté entre les espèces et permet donc de comprendre leur histoire évolutive. Cette nouvelle classification remet surtout en cause les taxons supérieurs, donc les critères qui sont très généraux.

1.2 Identification botanique des Angiospermes

L'identification des espèces végétales repose sur une classification pré-existante et donc diffère de la taxinomie car elle ne cherche pas à regrouper des espèces en taxon, mais à assigner l'individu observé à un groupe de la classification. Une fois que toutes les espèces sont rangées et les taxons ordonnés, il faut une méthode permettant de naviguer dans la hiérarchie de la classification pour assigner un taxon à une plante inconnue. En biologie, une clef de détermination est un outil d'identification de taxon qui consiste en un ensemble de choix à réaliser par l'observation de critères botaniques. Par une série d'options opposées, elle permet de cibler le groupe de végétaux auquel appartient l'échantillon à identifier, ceci sur la base des attributs de cet échantillon.

1.2.1 Clefs de détermination botaniques

Les premières clefs de détermination botaniques sont détaillées dans des ouvrages dénommées flores. Parmi les plus célèbres, nous pouvons citer la flore de Bonnier [Bon08], la flore de Coste [Cos03], ou la flore de Rouy [Rou13]. Cependant ces clefs utilisent préférentiellement les organes reproducteurs, donc principalement les fleurs. Ces premières clefs demandent à l'utilisateur d'analyser les différentes parties de la fleur (position de l'ovaire, nombre d'étamines, type d'inflorescence, *etc.*), souvent avec du vocabulaire botanique technique et demandant régulièrement une dissection de l'organe, pour descendre dans la hiérarchie de taxons (Figure 1.6).

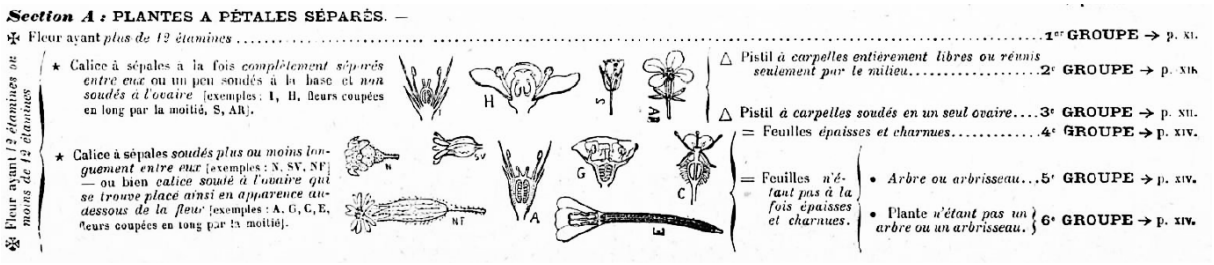


FIGURE 1.6 – Extrait de la flore de Bonnier [Bon08].

Ces ouvrages souvent anciens et austères, sont généralement perçus aujourd'hui comme destinés aux experts, et n'incitent pas forcément à l'apprentissage de la botanique du fait de leur vocabulaire complexe rendant difficile la navigation entre les clefs de détermination et de leur manque d'illustrations, comme nous pouvons le voir avec les premières clefs de détermination de la flore de Coste (Figure 1.7).

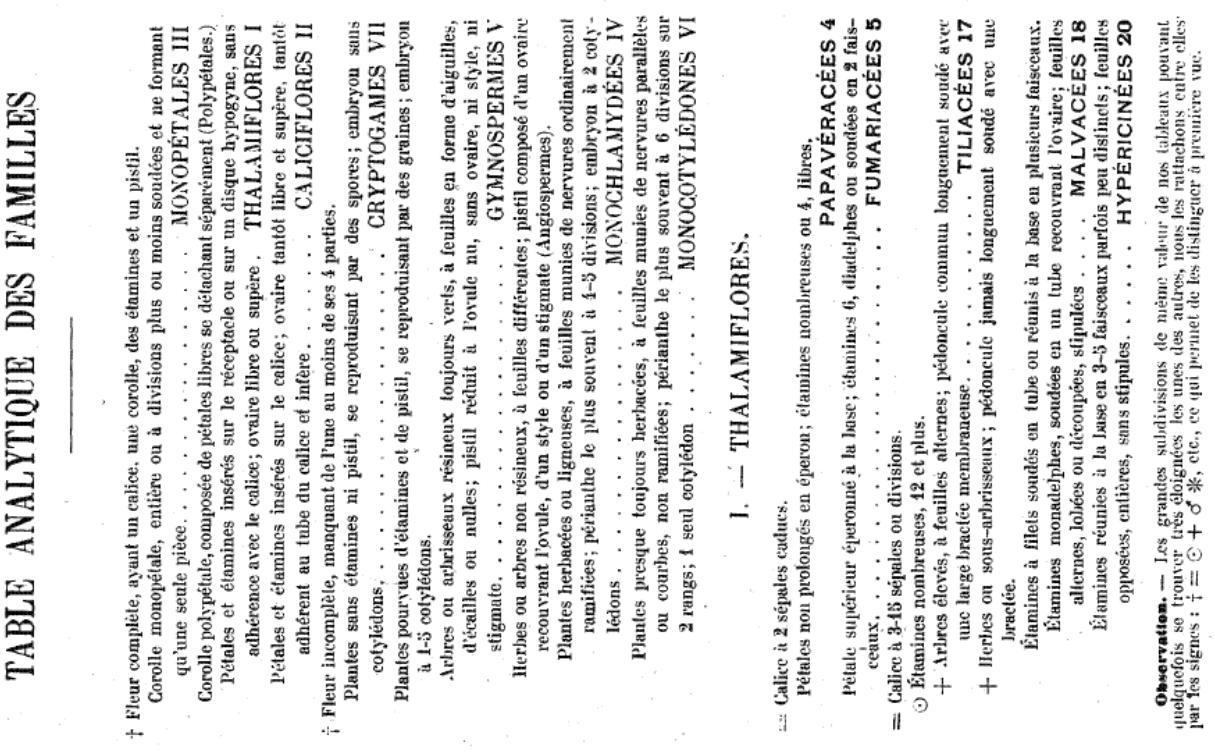


TABLE ANALYTIQUE DES FAMILLES

† Fleur complète, ayant un calice, une corolle, des étamines et un pistil.
Corolle monopétale, entière ou à divisions plus ou moins soudées et ne formant qu'une seule pièce. . . . . MONOPÉTALES III
Corolle polypétale, composée de pétales libres se détachant séparément (Polypétales.)
Pétales et étamines insérés sur le réceptacle ou sur un disque hypogyne, sans adhérence avec le calice; ovaire libre ou supère. THALAMIFLORES I
Pétales et étamines insérés sur le calice; ovaire tantôt libre et supère, tantôt adhérent au tube du calice et infère. . . . . CALICIFLORES II
†‡ Fleur incomplète, manquant de l'une au moins de ses 4 parties.
Plantes sans étamines ni pistil, se reproduisant par des spores; embryon sans cotylédons. . . . . CRYPTOGAMES VII
Plantes pourvues d'étamines et de pistil, se reproduisant par des graines; embryon à 1-3 cotylédons.
Arbres ou arbrisseaux résineux toujours verts, à feuilles en forme d'aiguilles, d'écailles ou nulles; pistil réduit à l'ovule nu, sans ovaire, ni style, ni stigmate. . . . . GYMNOSPERMES V
Herbes ou arbres non résineux, à feuilles différentes; pistil composé d'un ovaire recouvrant l'ovule, d'un style ou d'un stigmate (Angiospermes).
Plantes herbacées ou ligneuses, à feuilles munies de nervures ordinairement ramifiées; périanthe le plus souvent à 4-5 divisions; embryon à 2 cotylédons . . . . . MONOCHLAMYDÉES IV
Plantes presque toujours herbacées, à feuilles munies de nervures parallèles ou courbes, non ramifiées; périanthe le plus souvent à 6 divisions sur 2 rangs; 1 seul cotylédon . . . . . MONOCOTYLÉDONES VI

I. — THALAMIFLORES.

= Calice à 2 sépales caducs.
Pétales non prolongés en éperon; étamines nombreuses ou 4, libres. PAPAVÉRACÉES 4
Pétale supérieur éperonné à la base; étamines 6, diadelphes ou soudées en 2 faisceaux. . . . . FUMARIACÉES 5
≡ Calice à 3-18 sépales ou divisions.
☉ Étamines nombreuses, 12 et plus.
+ Arbres élevés, à feuilles alternes; pédoncule commun longuement soudé avec une large bractée membraneuse. . . . . TILIACÉES 17
+ Herbes ou sous-arbrisseaux; pédoncule jamais longuement soudé avec une bractée.
Étamines à filets soudés en tube ou réunis à la base en plusieurs faisceaux. Étamines monadelphes, soudées en un tube recouvrant l'ovaire; feuilles alternes, lobées ou découpées, stipulées . . . . MALVACÉES 18
Étamines réunies à la base en 3-5 faisceaux parfois peu distincts; feuilles opposées, entières, sans stipules. . . . . HYPÉRICINÉES 20

Observation. — Les grandes subdivisions de même valeur de nos tableaux peuvent quelquefois se trouver très éloignées les unes des autres; nous les rattachons entre elles par les signes : † = ☉ + ♂ ✻, etc., ce qui permet de les distinguer à première vue.

Figure 1.7 – Extrait de la flore de Coste [Cos03].

Les fleurs sont utilisées pour identifier les espèces car, au-delà du fait que la notion d'espèces est définie par le caractère de reproductibilité, ces organes sont les plus constants visuellement. C'est aussi la partie de la plante qui, historiquement, s'est montrée la plus discriminante tout en étant comparable. En effet, elles présentent toujours les mêmes caractéristiques (symétrie, position des pistils, étamines, *etc.*), contrairement aux feuilles ou aux écorces, qui dépendent de l'individu et peuvent varier avec l'âge de la plante. Ainsi, si nous prenons le cas du bouleau verruqueux (*Betula pendula*), son écorce évolue au fur et à mesure que l'arbre grandit, passant d'une écorce lisse et blanche à une écorce grise et rugueuse, en fonction des conditions environnementales (Figure 1.8).

Figure 1.8 – Évolution de l'écorce de *Betula pendula* entre un arbre jeune (à gauche) et un arbre plus âgé (à droite)

Par ailleurs, sur les organes non reproducteurs, il peut y avoir une grande similitude inter espèces et inter genres (entre deux espèces ou deux genres), comme présenté sur la Figure 1.9.

FIGURE 1.9 – Similitude entre l'écorce de *Quercus robur* à gauche et de *Acer platanoides* à droite, et entre une feuille de *Cornus mas* à gauche et de *Viburnum tinus* à droite

C'est généralement une fois que le genre est identifié que les clefs de détermination concernent les autres organes comme les feuilles ou la tige (écorce). C'est aussi une fois le genre identifié que d'autres informations, telle que la nature du terrain (plaine, montagne, *etc*) sont prises en compte pour l'identification de l'espèce.

1.2.2 Identification via les fleurs

Chez les Angiospermes, la fleur constitue l'organe de reproduction et l'ensemble des enveloppes (pétales, sépales, *etc.*) qui l'entoure. C'est à la suite de la pollinisation que la fleur se transforme en fruit contenant des graines, structures qui contiennent et protègent l'embryon végétal.

Comme nous l'avons vu précédemment, les fleurs sont les premiers organes utilisés par les botanistes pour identifier les taxons. La morphologie des fleurs et la position des différents éléments qui la composent, ainsi que la disposition des fleurs sur la tige sont des éléments caractéristiques particulièrement étudiés par les experts. Les différents organes de la fleur sont représentés sur la Figure 1.10.

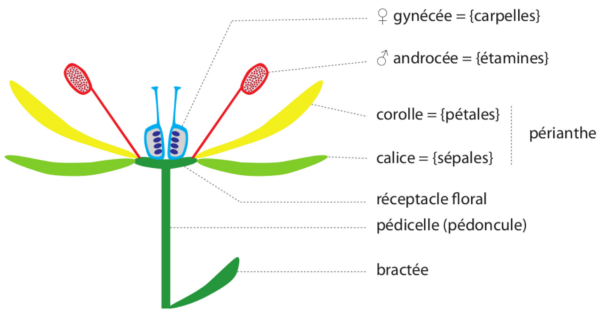


FIGURE 1.10 – Schéma d'une fleur par Hervé Sauquet.

Les fleurs peuvent porter les organes reproducteurs des deux sexes : les étamines (mâle) et les carpelles (femelle). L'ensemble des étamines constitue l'androcée, et l'ensemble des carpelles constitue le gynécée (ou pistil). Dans ce cas, la fleur est dite hermaphrodite. Dans le cas contraire la fleur est dite unisexuée. Le même individu d'une espèce peut porter les fleurs d'un seul sexe, c'est une plante unisexuée (par exemple *Actinidia deliciosa*, le plant de kiwi). Si une plante présente les fleurs des deux sexes, c'est une plante bisexuée, qui peut avoir des fleurs mâles et des fleurs femelles sur le même individu (plante monoïque, par exemple *Castanea sativa*, le châtaigner, voir Figure 1.11), ou des fleurs hermaphrodites (par exemple *Nerium oleander*, le laurier-rose).

Figure 1.11 – Fleur femelle et fleurs mâles du châtaignier (*Castanea sativa*).

Les fleurs sont rarement solitaires sur une plante. Elles sont regroupées en inflorescences, ce qui permet une augmentation de l'attraction des pollinisateurs. Un type d'inflorescences (Figure 1.12) correspond à une disposition particulière des fleurs les unes par rapport aux autres. Le type d'inflorescence est souvent caractéristique du taxon famille, ce pourquoi elle est fréquemment exploitée. Les inflorescences sont très variées et peuvent être combinées, par exemple en grappe de grappes (une panicule), en corymbe de capitules, *etc.*). Sans une observation attentive, il est difficile de faire la différence entre certaines inflorescences.

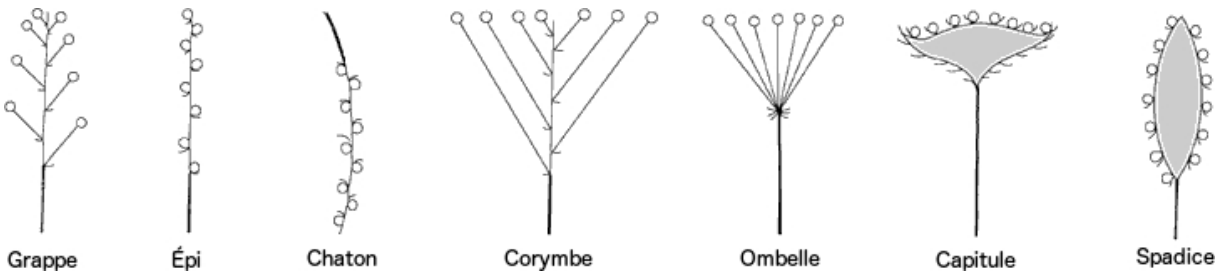


Figure 1.12 – Schéma des différents types d'inflorescences.

Un autre élément employé pour la reconnaissance de végétaux par les fleurs est la symétrie de celles-ci. Les éléments floraux s'organisent plus ou moins régulièrement autour de l'axe floral (Figure 1.13). Il existe deux types de symétrie florale :

— les fleurs actinomorphes (aktinos = rayon), ou régulières, ont un centre de symétrie ;
— les fleurs zygomorphes (zygos = couple), ou irrégulières, ont une symétrie bilatérale.

FIGURE 1.13 – Illustration des deux types de symétrie chez les fleurs de *Prunus avium* (actinomorphe) et de *Cercis siliquastrum* (zygomorphe).

Pour résumer la morphologie de la fleur, le botaniste américain John Henry Schaffner (1866-1939) a proposé le diagramme floral [Sch16]. Il s'agit d'une représentation schématique de l'organisation des différentes pièces florales qui permet d'identifier facilement la famille d'une plante et de simplifier la compréhension de l'organisation la fleur. On représente les différentes pièces florales en position anatomique comme pour une coupe transversale de la fleur de l'extérieur vers l'intérieur. En complément du diagramme floral, la formule florale permet de représenter la structure d'une fleur à l'aide de chiffres, de lettres et de symboles divers, en présentant les informations importantes à propos de la fleur sous une forme compacte. La formule florale peut servir à représenter une espèce particulière, mais elle peut être généralisée pour caractériser des taxons de niveau plus élevé, en indiquant des plages pour les nombres de pièces florales (Figure 1.14).

FIGURE 1.14 – Fleur de pomme de terre (*Solanum tuberosum*) et diagramme floral dont la formule florale est K : (5) ; [C : (5) ; A : 5] ; Gsup : (2)

Lorsqu'un botaniste est face à une plante à fleurs inconnue, il va rarement s'intéresser à la couleur de la fleur. En effet, la couleur de la fleur n'est pas caractéristique des taxons (famille ou genre). De plus, pour une même espèce, la couleur des fleurs peut varier énormément en fonction du cultivar. Un cultivar est une variété de plante obtenue en culture, généralement par sélection, pour ses caractéristiques. Nous pouvons prendre le cas des lauriers roses (*Nerium oleander*, voir Figure 1.15).

Figure 1.15 – Fleurs de deux individus de *Nerium oleander* de cultivars différents.

Cependant pour l'étude des cultivars, la couleur des fleurs a été codifiée pour la reconnaissance et la création de variétés de fleurs, notamment pour les roses. La Royal Horticultural Society (RHS) a ainsi développé un dictionnaire de couleur. Les blocs de couleurs de la charte de la RHS sont organisés ainsi :

— rouge à rose
— bleu à violet
— orange à jaune
— blanc à crème
— vert à brunâtre

Dans ces cas-là, c'est uniquement la couleur de la corolle (les pétales) qui est utilisée, et non celle du centre de la fleur. Il existe des plantes à fleurs qui ont des pétales de couleurs différentes sur une même corolle, ou avec des motifs. C'est alors la couleur dominante de la fleur qui est utilisée.

1.2.3 Identification via les fruits

Après la fécondation, la fleur se transforme en fruit qui permet de protéger les graines et d'aider à leur dissémination. Une fois le fruit mûr, il se détache de l'arbre ou s'ouvre (organe déhiscent). Ce sont les parois de l'ovaire qui se transforment lors de la fructification. On peut y distinguer trois enveloppes dont l'épaisseur et la texture vont varier selon la nature du fruit. L'ensemble des trois enveloppes est appelé péricarpe (Figure 1.16). Dans certains cas, d'autres parties de la plante, généralement le réceptacle floral, participent à la formation du fruit.

L'organisation du, ou des, ovaires transformés au sein du fruit est détaillée par un vocabulaire botanique présenté dans la Table 1.1.

De même que pour les fleurs et les inflorescences, il est difficile de faire la distinction entre les différents types de fruits sans dissection, d'autant plus pour un novice.

Bien que directement liés aux fleurs, les fruits ne sont que très peu utilisés par les botanistes pour l'identification des espèces de végétaux. En effet, le fruit est un organe vivant qui passe par différents aspects avant sa maturité. Beaucoup de plantes comptent

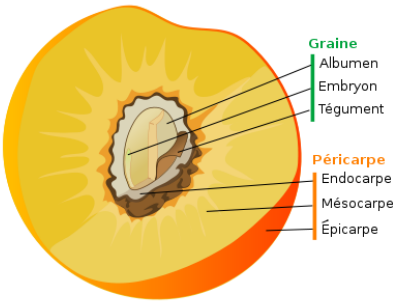


FIGURE 1.16 – Structure schématique d'un fruit typique.

Simples	Secs	Secs indéhiscents	Akène : Fruit ne s'ouvrant pas, à une seule graine non soudée aux parois du fruit (noisette).
			Caryopse : Fruit ne s'ouvrant pas, à une seule graine soudée aux parois du fruit (céréales).
		Secs déhiscents	Capsule : Fruit s'ouvrant par un couvercle ou par des pores.
			Gousse : 1 carpelle à 2 fentes de déhiscence.
			Follicule : 1 carpelle à 1 fente de déhiscence.
			Silique : 2 carpelles avec une fausse cloison.
	Charnus	Drupes	Une seule graine (rarement plusieurs) protégée par un noyau (Abricot, cerise, olive).
		Baies	Plusieurs graines (rarement une seule) directement dans la pulpe (Raisin, groseille).
Multiples	Polyakènes		Ensemble d'akènes (Renoncule, potentille).
	Drupéoles		Ensemble de drupes (Framboise)
Comlexes			Fruits (au sens gastronomique) formés d'une partie de réceptacle, ou de gynécée, et englobant les vrais fruits (Fraise, pomme).

TABLE 1.1 – Description des différents fruits et de leurs propriétés morphologiques.

sur les animaux pour disséminer leurs graines (zoochorie), plus particulièrement en les ingérant (endozoochorie). Les fruits changent de couleur au cours de leur existence, prenant dans la plupart des cas des couleurs vives pour attirer les animaux. Pour certaines espèces végétales, la formation du fruit est assez longue, ne permettant pas une description unique (voir Figure 1.17).

FIGURE 1.17 – Développement du fruit (gland) chez *Quercus pubescens*.

1.2.4 Identification via les feuilles

La feuille est un organe essentiel de la plante, car c'est celui dédié à la photosynthèse et permettant des échanges entre la plante et son environnement (respiration et transpiration). Chez les Angiospermes, comprenant les arbres qualifiés de "feuillus", les feuilles sont généralement plates et présentent une très grande variabilité de forme, contrairement aux Gymnospermes qui regroupent les conifères portant des aiguilles. Malgré sa grande variabilité de forme, la feuille n'est pas un organe prépondérant dans l'identification végétale. En effet, les traits d'une feuille ne sont pas caractéristiques d'un taxon. De plus, certaines espèces végétales peuvent présenter des feuilles de formes différentes bien qu'appartenant à une même espèce, voir à un même individu (on parle alors d'hétérophyllie voir Figure 1.18).

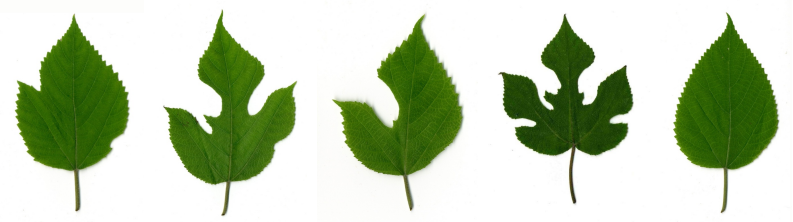


Figure 1.18 – Variabilité intra-classe des feuilles au sein du mûrier à papier (*Broussonetia papyrifera*).

Cependant, leurs caractéristiques morphologiques sont bien définies par un vocabulaire botanique et il est utilisé pour différencier les espèces présentes au sein d'un même genre. La structure de la feuille chez les Angiospermes est présentée sur la Figure 1.19. Le limbe est la partie plate principale de la feuille qui contient les cellules permettant la photosynthèse. Il est relié à la tige de la plante par le pétiole. Il se peut que la feuille soit directement attachée à la tige chez certaines espèces végétales. Dans ce cas les feuilles sont dites sessiles. L'apex est le sommet (la pointe) du limbe, et la base la partie opposée où s'attache le pétiole. Les nervures sont le prolongement des tissus vasculaires issus du pétiole dans le limbe. C'est un ensemble de vaisseaux conduisant la sève dans la feuille, dont leur disposition forme la nervation.

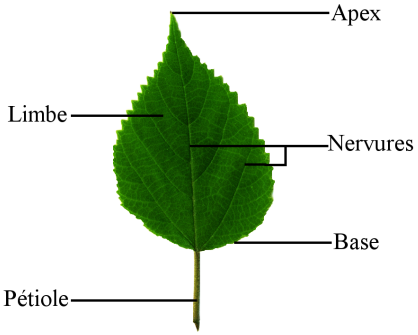


Figure 1.19 – Schéma général d'une feuille d'Angiosperme.

Les morphologies des différentes parties de la feuille sont des critères de description botanique très bien détaillées. La feuille est en premier caractérisée par sa forme générale, particulièrement si elle présente des lobes, c'est-à-dire une division du limbe. Les différents types sont présentés dans le Tableau 1.2.

TABLE 1.2 – Les différentes formes générales de feuille. Issue de [Cer13b].

	Simple	limbe continu, sans séparation apparente
	Palmatilobée	limbe séparé en plusieurs lobes aux nervures se rejoignant à l'extrémité du pétiole
	Pennatilobée	limbe séparé en plusieurs lobes aux nervures échelonnées sur la nervure principale

Pour le cas des feuilles simples, la forme du limbe est décrite selon un vocabulaire plutôt géométrique dont les modèles sont détaillés dans le Tableau 1.3.

TABLE 1.3 – Les différentes formes de feuille simple. Issue de [Cer13b].

	Linéaire	très étroite et allongée, aux bords quasiment parallèles
	Lancéolée	étroite et allongée, pointue au sommet et généralement plus large à la base
	Elliptique	terme vague, étroite ou large, généralement arrondie et égale à la base et au sommet
	Oblongue	allongée, aux bords quasiment parallèles, arrondie à la base et au sommet
	Ovale	en forme d'œuf, de largeur moyenne, plus large et arrondie à la base qu'au sommet
	Obovale	opposé à ovale, de largeur moyenne, plus large et arrondie au sommet qu'à la base
	Rhomboïdale	de la forme d'un losange, de largeur moyenne, égale à la base et au sommet
	Spatulée	assez étroite, plus large et arrondie au sommet, la base s'étirant vers le pétiole
	Orbiculaire	de la forme d'un cercle, très large et très arrondie à la base et au sommet
	Suborbiculaire	presque orbiculaire, un peu moins large que longue mais très arrondie
	Triangulaire	de la forme d'un triangle, plus ou moins large à la base, très pointue au sommet
	Cordée	de la forme d'un cœur, large et incurvée à la base, plus ou moins pointue au sommet

Il faut cependant noter que les limites entre ces différentes classes sont floues, et que selon les conditions dans lesquelles pousse la plante, une même espèce peut avoir des feuilles dont le limbe est plus ou moins large, présentant alors des feuilles de forme linéaire, lancéolée ou oblongue.

Le limbe porte aussi une autre information pouvant qualifier l'espèce végétale, qui est la nervation. La structure et l'agencement des différentes nervures sont aussi classifiés en plusieurs types détaillés dans le Tableau 1.4. Pour déterminer le type de nervation d'une

feuille, il est plus aisé de regarder le dessous de la feuille, c'est-à-dire la face qui n'est pas orientée vers le soleil.

Table 1.4 – Les différentes formes de nervures. Issue de [Cer13b].

	Uninervée	une seule nervure principale suivant l'axe principal de la feuille
	Parallèle	plusieurs nervures orientées selon l'axe principal de la feuille sans intersection
	Pennée	plusieurs paires de nervures secondaires échelonnées le long de la nervure principale
	Alterne	plusieurs paires de nervures secondaires mais pas en face les unes des autres
	Arquée	plusieurs paires de nervures secondaires se rejoignant vers l'apex de la feuille
	Palmée	un nombre impair de nervures principales divergeant depuis la base de la feuille
	Réticulée	nervures secondaires et tertiaires formant un réseau complexe

De même que pour la forme générale, il existe un vocabulaire spécifiant les formes de base (Table 1.5) et d'apex (Table 1.6). Ici aussi, il peut être complexe de ne pas confondre certaines classes d'apex ou de base. Également, la feuille d'une espèce ne correspondra pas à un seul type d'apex ou de base, mais peut varier. Par exemple, les feuilles d'olivier (*Olea europaea*) peuvent présenter des apex de forme aiguë à obtuse.

Table 1.5 – Les différentes formes de bases de feuille. Issue de [Cer13b].

	Atténuée	limbe très étroit à la base, qui se prolonge en pétiole
	En coin	limbe assez étroit à la base, qui forme un angle aigu assez marqué
	Arrondie	limbe plutôt large à la base, qui forme un angle obtus arrondi et peu marqué
	Oblique	limbe à la base dissymétrique, l'angle s'ouvrant plus largement d'un côté que de l'autre
	Décrochée	limbe à la base dissymétrique, un côté du limbe s'attachant plus bas que l'autre
	Tronquée	limbe aplati à la base, qui forme un angle plat sur une largeur plus ou moins importante
	En Cœur	limbe large à la base, qui forme un angle rentrant, qui s'arrondit de part et d'autre
	Echancrée	limbe arrondi à la base, qui forme une encoche plus ou moins profonde
	Auriculée	limbe assez large à la base, qui forme deux petites bosses de part et d'autre

Le dernier aspect de la feuille qui est décrit par les botanistes est le bord du limbe, appelé également la marge. L'aspect de la marge dépend du fait que les nervures touchent le bord du limbe ou non. Les différents types de marge sont décrits dans le Tableau 1.7. La confusion entre les différents types de marge vient de la taille des différentes dents et de leurs angularités.

TABLE 1.6 – Les différentes formes d'apex de feuille. Issue de [Cer13b].

	Acuminée	limbe très étroit au sommet, qui se termine par une pointe effilée plus ou moins longue
	Aiguë	limbe se rétrécissant au sommet, en formant une pointe avec un angle aigu
	Obtuse	limbe se rétrécissant au sommet, en formant une pointe avec un angle obtus
	Epineuse	limbe se rétrécissant au sommet, qui se termine par une pointe voire une épine
	Recourbée	limbe se rétrécissant au sommet, qui se termine par une pointe inclinée d'un coté
	Cuspidée	limbe se rétrécissant au sommet, qui se termine par une pointe courte marquée et rigide
	Mucronée	limbe se rétrécissant au sommet, qui se termine par une pointe courte dure et arrondie
	Arrondie	limbe assez large au sommet, sans pointe mais s'arrondissant
	Tronquée	limbe aplati au sommet, qui forme un angle plat plus ou moins large
	Rétuse	limbe assez large au sommet, qui forme un angle rentrant
	Echancrée	limbe assez large au sommet, qui forme une encoche marquée

TABLE 1.7 – Les différentes formes de bord de feuille. Issue de [Cer13b].

	Entière	bord parfaitement lisse, sans aucune aspérité ni aucune échancrure
	Ciliée	bord présentant à sa marge de petits poils ou cils fins
	Ondulée	bord sans aspérités mais présentant des sinuosités larges et peu marquées
	Crénelée	bord présentant des dents de petite taille et très arrondies
	Denticulée	bord présentant des dents marquées, mais très fines
	Dentée	bord présentant des dents bien marquées, qu'elles soient de taille petite ou moyenne
	Dentée en scie	bord présentant des dents bien marquées, toutes orientées vers l'apex de la feuille
	Doublement dentée	bord présentant deux niveaux de dents, les petites entre les plus grandes
	Epineuse	bord présentant dents larges et très pointues, voire des épines
	Anguleuse	bord sans dents mais présentant angles larges et assez marqués
	Sinuée	bord sans dents mais présentant des sinuosités marquées, de taille moyenne
	Lobée	bord présentant des lobes plus ou moins marqués, et de larges sinus
	Pennatifide	bord présentant des sinus larges et profonds, presque jusqu'à la nervure

Nous pouvons voir qu'il existe un vocabulaire botanique très riche pour caractériser la forme d'une feuille. Ainsi, associer une forme de limbe, de base ou d'apex à une feuille, définie par un vocabulaire spécifique, est une tâche complexe pour un novice en botanique. Contrairement aux fleurs, la couleur de la feuille n'est pas vraiment utilisée, car celle-ci change au cours de l'année (spécialement en automne). Elle est précisée quand la couleur de la face supérieure est différente de la face intérieure, ou si la feuille a un aspect luisant.

Jusqu'à maintenant, nous avons vu uniquement les feuilles dites simples. dont le limbe est entier. Or, il se peut que le limbe de la feuille soit découpé en plusieurs petites feuilles appelées folioles. Dans ce cas, les feuilles sont dénommées feuilles composées. La disposition des folioles le long du rachis peut être caractéristique d'une espèce. Les différentes dispositions sont présentées dans le Tableau 1.8.

Table 1.8 – Les différents de feuilles composées. Issue de [Cer13b].

	Simple	limbe d'un seul tenant, continu
	Trifoliée	limbe divisé en trois folioles uniquement qui se rejoignent à l'extrémité du pétiole
	Digitée	limbe divisé en plusieurs folioles qui se rejoignent toutes à l'extrémité du pétiole
	Paripennée	limbe divisé en un nombre pair de folioles échelonnées le long du rachis
	Imparipennée	limbe divisé en un nombre impair de folioles avec une foliole terminale
	Bipennée	limbe divisé en folioles disposées le long du rachis et elles-mêmes pennées

La différence entre ces types de feuilles composées est plus abordable pour des novices en botanique que la caractérisation de la forme du limbe. Les folioles sont ensuite caractérisés de la même manière que le limbe des feuilles simples comme vu précédemment.

Pour finir, et de manière plus simple, c'est la position des feuilles le long de la tige de la plante (la phyllotaxie) qui peut apporter une information sur l'espèce végétale (Table 1.9).

Table 1.9 – Les différentes positions d'attache des feuilles à la tige.

	Alternées	Folioles rattachées en alternance à différents niveaux
	Opposées	Paires de folioles rattachées l'une en face de l'autre aux mêmes niveaux
	Verticillées	Au moins trois folioles rattachées en cercle à un même niveau
	Rosette	Feuilles étalées en cercle au niveau du collet (zone de transition entre la racine d'une plante et la surface)

1.2.5 Identification via les écorces

L'organe le moins utilisé par les botanistes pour l'identification des espèces végétales est sans doute la tige pour les plantes en possédant une ou plusieurs (dans le cas contraire

on parle de plante acaule). Habituellement cylindrique, la forme de la tige peut être carrée, trigone, renflée, noueuse, striée, cannelée *etc.* D'autres caractères, comme la pilosité et les excroissances peuvent être des caractéristiques plus stables et peuvent plus facilement être utilisés pour la détermination de l'espèce. La surface d'une tige peut être lisse, pubescente (présentant des poils fins), épineuse, verruqueuse, ou subéreuse (qui contient du liège).

Un cas particulier de tige est le tronc pour le cas des arbres et arbustes. D'un point de vue botanique, un arbre ou un arbuste est une plante ligneuse terrestre, qui se caractérise par la croissance secondaire en épaisseur de son tronc permettant un développement en hauteur et une ramification des branches. C'est-à-dire que la plante pousse d'abord vers le haut (croissance primaire) avant de croître en diamètre. Une plante ligneuse est une plante qui produit de la lignine, une molécule de cellulose constituant le bois. La notion de plante ligneuse n'est pas classificatoire c'est-à-dire qu'elle ne fournit pas d'information sur le taxon. Le tronc, et plus particulièrement l'écorce, n'est pas non plus un organe très étudié pour l'identification végétale. En effet, la structure de l'écorce dépend de nombreux paramètres tels que la situation géographique, l'altitude, l'âge, la nature du sol, les conditions hydrologiques, l'ensoleillement, *etc.* Mais le facteur majoritaire est l'âge du sujet. Il est quasiment impossible d'identifier un arbre dans ses jeunes années à partir de son écorce, contrairement à ses fleurs ou ses feuilles. Chez la plupart des arbres, le phellogène (tissu végétal qui produit le suber, ensemble de cellules allant vers l'extérieur du tronc) ne peut pas compenser la croissance en diamètre par un étirement suffisant. Les cellules situées en dehors du phellogène meurent et forment l'écorce qui, sous la poussée des autres cellules se fissure [God12].

Nous pouvons cependant citer le travail de Michael Wojtech qui a rédigé un livre [Woj11] sur la reconnaissance des espèces d'arbres de la côte nord-est d'Amérique, et introduit différentes familles pour classer les écorces en fonction de leur structure d'un point de vue visuel. Les différentes familles sont présentées sur la Figure 1.20.

Après avoir identifié le type auquel appartient l'écorce, les botanistes prennent en considération la couleur et d'autres éléments structurels tels que l'organisation des excroissances entre elles ou la forme des lenticelles. Les lenticelles sont l'équivalent des pores. Elles sont présentes sur le liège à la surface de l'écorce des arbres et forment des aspérités colorées pour permettre des échanges gazeux avec l'extérieur (Figure 1.21).

Pour finir, il faut noter que l'identification botanique d'une plante est d'autant plus compliquée que le nombre d'espèces possibles est grand (près de 800 genres présents dans la flore de Bonnier). Dans le cadre du projet ReVeRIES, nous nous restreignons à l'identification des espèces d'arbres et arbustes de plus de 1m20 présents en France métropolitaine (et parlerons d'identification d'espèces d'arbres pour simplifier).

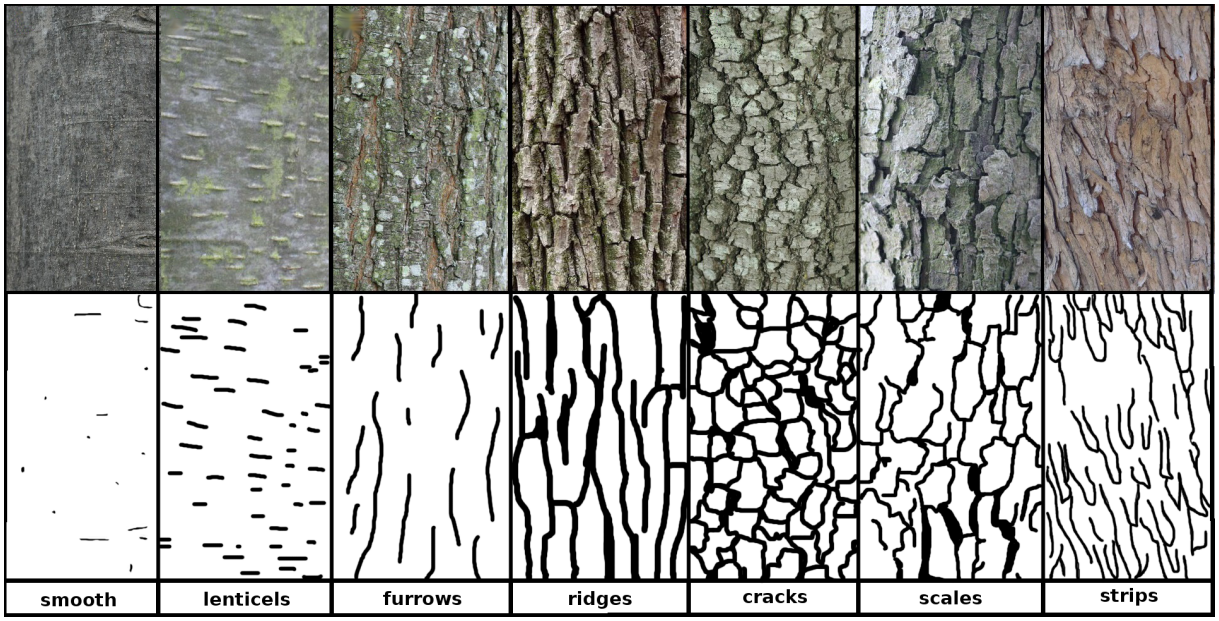


Figure 1.20 – Différentes familles de texture d'écorce selon Michael Wojtech [Woj11] avec mise en évidence des structures (dessin).

Figure 1.21 – Exemple de lenticelles de différentes formes : rondes, allongées, losangiques.

1.3 Reconnaissance de végétaux par des novices

Nous avons vu qu'il existe des outils et un vocabulaire associé à destination des botanistes pour permettre l'identification des végétaux. Si nous mettons de côté la difficulté pour un novice en botanique d'identifier des morphologies végétales à partir d'un vocabulaire technique, nous nous apercevons que la frontière entre une fleur et un fruit est incertaine. En effet, lorsque nous parlons de fruits, nous avons tendance à penser aux fruits que nous consommons, c'est-à-dire les pommes, les pêches, les bananes, les raisins, *etc.* Or sur les arbres d'ornement, il existe une très grande variété de forme de fruit. L'utilisateur novice peut donc confondre les fleurs, les fruits et les bourgeons (voir Figure 1.22).

Ainsi, pour un novice en botanique, un fruit est tout ce qui n'est pas une feuille ou une fleur « classique » (colorée et avec des pétales). Pour compliquer la distinction entre les

(a) (b) (c) (d)

FIGURE 1.22 – Les fruits de de *Broussonetia papyrifera* a) et d'*Euonymus europaeus* b) sont confondus avec des fleurs à cause de leur aspect très coloré. Les chatons de *Carpinus betulus* c) et les bourgeons de fleurs de *Viburnum tinus* d) sont confondus avec des fruits.

fruits et les fleurs, certaines espèces ont leurs fleurs qui se transforment progressivement en fruit sans que leur morphologie ne soit significativement modifiée. Nous pouvons prendre le cas du platane (*Platanus* x *hispanica*) ou du bouleau (*Betula pendula*), où il y a un léger changement d'aspect ou de couleur entre la fleur et le fruit (Figure 1.23).

(a) *Platanus x hispanica* (b) *Betula pendula*

FIGURE 1.23 – Fleur (à gauche) et fruit (à droite) de *Platanus* x *hispanica*. Fleur (à gauche) et fruit (à droite) de *Betula pendula*.

Selon Valéry Malécot, enseignant-chercheur botaniste impliqué dans le projet ReVeRIES et appartenant au laboratoire IRHS, un utilisateur débutant en botanique va identifier comme fruit ou fleur tout organe sur l'arbre qui n'est pas une feuille. Pour confirmer cette hypothèse, notamment avec les organes des arbres d'ornement non cultivés pour leurs fruits, un questionnaire a été créé pour savoir comment des personnes qui ne sont pas expertes en botanique identifient une partie de la plante qui n'est pas une feuille. Plus précisément, nous cherchons à voir si ceux-ci sont capables de faire la différence entre des fleurs ou des fruits.

1.3.1 Questionnaire de différenciation des fruits et des fleurs

Le questionnaire est composé de 40 images d'organes d'arbres ou arbustes, soit des fruits, soit des fleurs. Pour chaque image, l'utilisateur doit indiquer s'il pense que la partie végétale qui ne fait pas partie du feuillage est une fleur, ou un fruit, ou autre. Dans le questionnaire il n'y a actuellement que des images de fruits ou de fleurs. La fleur ou le

fruit n'est pas segmenté/isolé du fond, pour mimer ce qu'un utilisateur peut rencontrer dans la nature. Pour chaque image, nous proposons cependant le choix de cocher "Autre". En effet, le but du questionnaire est aussi de déterminer ce qu'un utilisateur d'une application d'aide à l'identification de végétaux va prendre en photo si celle-ci lui demande de photographier une fleur ou un fruit. S'il n'identifie pas un organe comme fleur ou fruit (donc comme organe "Autre"), il ne va pas le prendre en photo, et indiquera à l'application qu'aucun des deux organes n'est présent sur l'arbre. D'autre part, il n'y a pas de choix "Je ne sais pas" pour "forcer" l'utilisateur à choisir entre fleur ou fruit. Chaque image est labellisée avec le numéro de la question, et non le nom de l'espèce, pour éviter que les personnes recherchent des informations sur la plante. Un exemple de question est présenté sur la Figure 1.24.

Figure 1.24 – Exemple de question pour le formulaire de distinction entre les fruits et les fleurs.

Le questionnaire a été envoyé à tous les membres du laboratoire LIRIS, ne travaillant donc pas dans le domaine de la botanique, ainsi qu'à l'ensemble des chercheurs du projet ReVeRIES. Avant de commencer à identifier les organes comme des fleurs ou des fruits, il est demandé à l'utilisateur de renseigner son niveau en botanique/reconnaissance des plantes. En tout, 128 personnes ont répondu au questionnaire. La Figure 1.25 présente la répartition du niveau botanique des participants selon leur propre évaluation.

Nous pouvons voir que, à quelques exceptions, les personnes questionnées ne sont pas expertes en botanique et correspondent au type de population qui pourrait utiliser une application d'identification botanique.

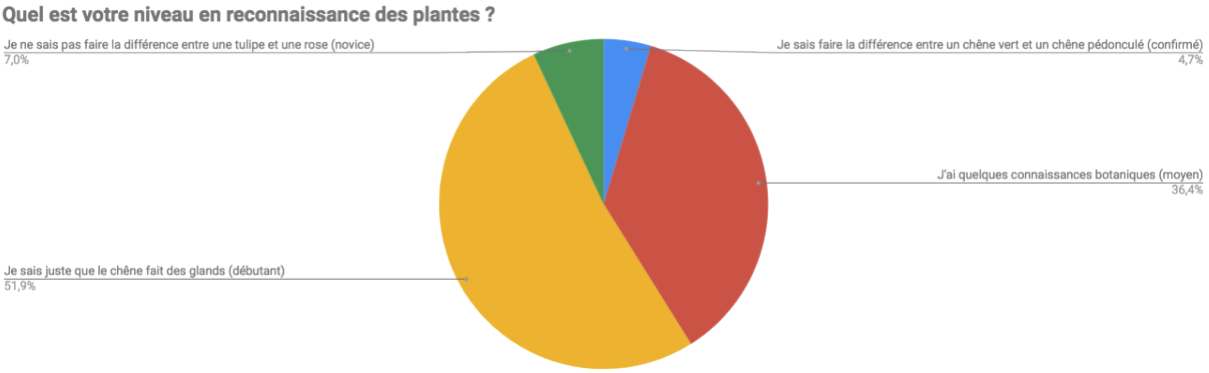


FIGURE 1.25 – Résultats pour l'auto-évaluation du niveau botanique de l'utilisateur.

Sur les résultats des 40 photos, nous avons trois catégories : les fleurs ou les fruits bien identifiés (Figure 1.26), les fruits ou fleurs mal identifiés (Figure 1.27), et les fruits ou fleurs bien identifiés mais avec une grande incertitude (Figure 1.28).

FIGURE 1.26 – Exemple de fleurs et de fruits bien reconnus par les utilisateurs.

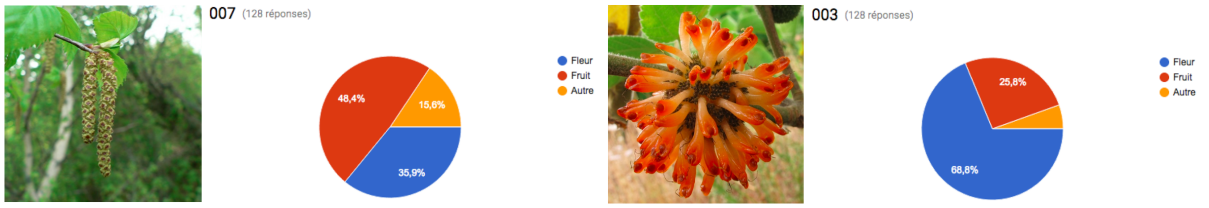


FIGURE 1.27 – Option fausse préférentiellement choisie entre le fruit et la fleur.

À la suite de ce questionnaire, nous avons eu la confirmation que des utilisateurs non experts en botanique ne savent généralement pas différencier une fleur d'un fruit d'un arbre au sens botanique du terme. Au vu des résultats, il apparaît qu'un organe est identifié de manière quasi certaine comme une fleur si celui-ci est coloré ou présente de manière distinctive des pétales (une corolle) ou une forme de fleur présente chez un

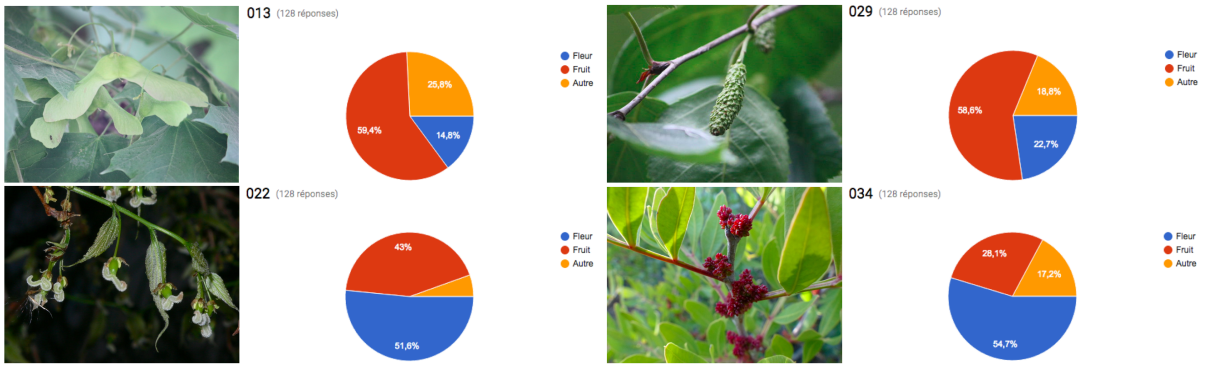


Figure 1.28 – Option correcte préférentiellement choisie, mais grande incertitude.

fleuriste. Ainsi, l'organe 003 (Figure 1.27) qui est le fruit du murier à papier (*Broussonetia papyrifera*) a été classifié par la majorité des personnes interrogées comme une fleur de part son aspect très coloré. Les fruits bien classifiés comme tels sont des fruits charnus (baies, drupes) ou secs, semblables à ceux trouvés chez un maraîcher. D'autre part, nous pouvons voir une augmentation du nombre de réponse "Autre" pour les organes dont l'aspect ne se rapproche pas de ce qui est couramment rencontré par la population dans les magasins (Figure 1.28). Si l'utilisateur n'identifie pas un organe comme fleur ou fruit (donc comme organe "Autre") il ne va pas le prendre en photo et indiquera à l'application d'identification d'espèces végétales qu'aucun des deux organes n'est présent sur l'arbre. Autrement, l'utilisateur risque de prendre en photo un fruit mais indiquera à l'application qu'il s'agit d'une fleur, ou inversement, ce qui pourrait entraîner une mauvaise prédiction de l'espèce.

1.4 Utilisation de clefs de détermination pour novices

Nous avons vu dans les parties précédentes comment sont reconnues les espèces d'arbres par les botanistes via les caractéristiques morphologiques des différents organes. Les Flores sont des outils très efficaces pour associer un individu aux taxons, mais restent difficiles à utiliser pour des novices sans connaissances de l'organisation taxinomique des espèces végétales entre elles. D'autre part, le problème de l'utilisation des organes reproducteurs pour l'identification est qu'ils ne sont présents sur l'arbre que sur une période assez réduite de l'année. La période de floraison ou de fructification peut en revanche être un indice pour l'identification. L'étude des organes reproducteurs nécessite de connaître un vocabulaire et d'avoir des compétences techniques que le botaniste néophyte n'a pas, sans parler de la subtilité à classer un organe en fruit ou fleur. Il existe cependant de nombreux outils pour aider à l'identification des espèces végétales se basant sur les différents organes, que nous allons décrire dans cette partie.

1.4.1 Clefs "papier"

Au cours des dernières années, des avancées ont été faites pour rendre plus agréables les flores. C'est principalement le vocabulaire ainsi que le nombre d'illustration qui ont été modifiés. Le vocabulaire est rendu plus accessible en enlevant les mots techniques botaniques pour les remplacer par leurs définitions ou par des schémas. Par exemple, une feuille ne sera plus caractérisée comme "pubescente", mais comme "couverte d'un duvet ou de poils fins". Dans ces guides grand public une grande part de l'ouvrage est consacrée aux illustrations (Figure 1.29). Quasiment chaque espèce est illustrée d'une ou plusieurs images pour permettre au lecteur d'identifier l'espèce sans vocabulaire. Quand les fleurs sont prises en compte, les espèces sont classées selon la couleur des fleurs et non en fonction de la famille. Les fruits et les fleurs n'étant pas présents tout au long de l'année, les ouvrages de flore simplifiée présentent des clefs de détermination sur les feuilles.

Finalement, il n'est pas possible d'avoir un ouvrage regroupant tous les végétaux d'un pays avec uniquement une clef de détermination sur les feuilles. Ainsi les ouvrages destinés à un public non botanique présentent généralement une sélection réduite d'espèces (les plus communes) en s'arrêtant au taxon du genre. En conséquence ces ouvrages apprennent à faire la différence entre un tilleul (genre *Tilia*) et un chêne (genre *Quercus*), mais pas entre un chêne pédonculé (*Quercus robur*) et un chêne chevelu (*Quercus cerris*).

Les inconvénients principaux de ces ouvrages sont qu'ils sont assez encombrants, et que l'utilisateur doit fouiller tout le livre à la recherche de la bonne espèce.

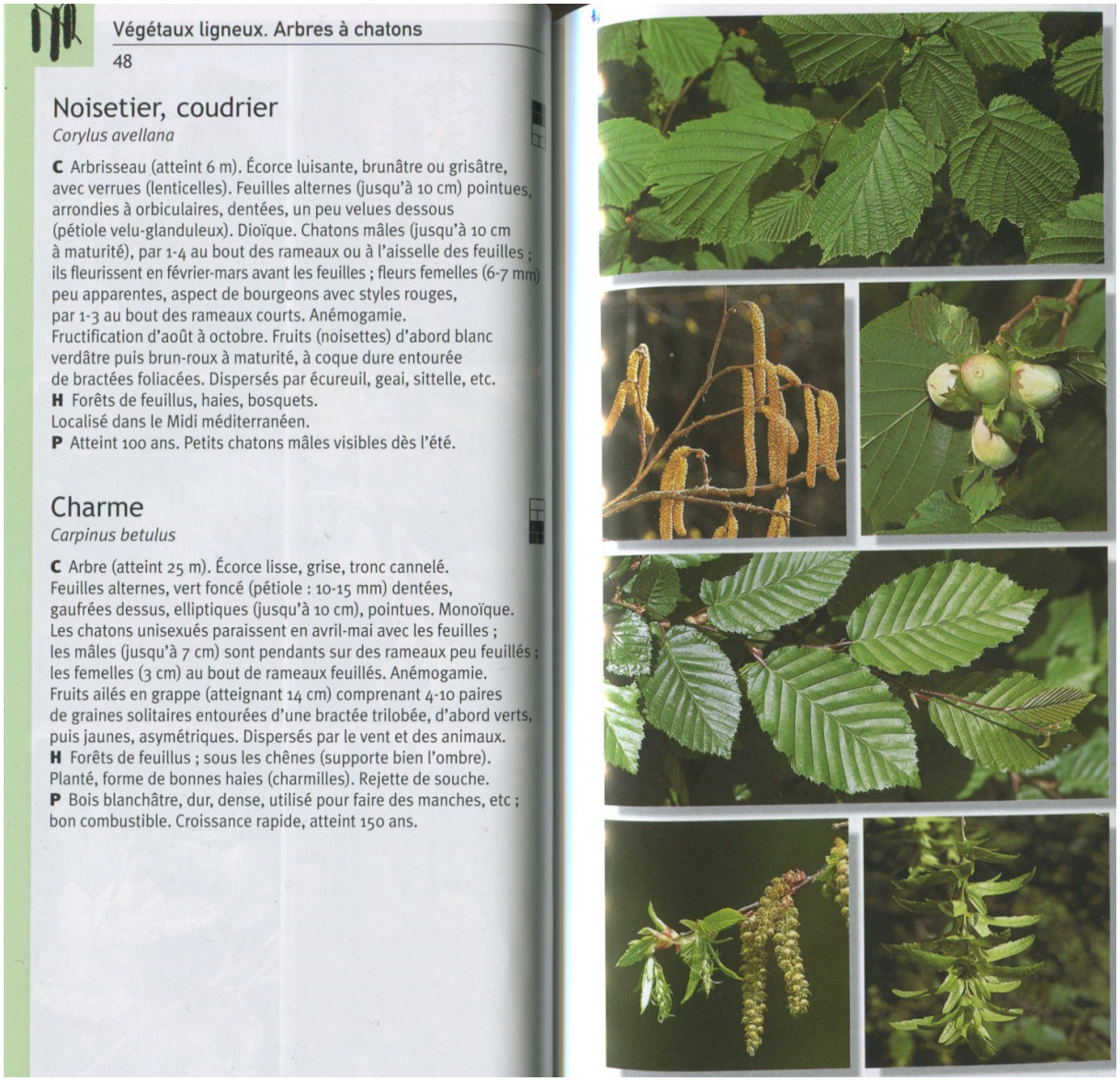
1.4.2 Clefs numériques

Avec l'avènement du numérique, ces flores ont été numérisées, et à chaque clef de détermination l'utilisateur n'a plus à feuilleter l'ouvrage pour trouver la clef suivante. Ainsi la Flore de Bonnier [Bon08] a été informatisée [1]. L'utilisateur doit cliquer sur les différents choix des questions (clefs) présentes dans l'ouvrage. Nous pouvons citer ID-Botanica [2] qui permet aussi de sauter certaines étapes dans la clef de détermination ; étapes pour lesquelles l'utilisateur n'a pas la réponse ou a du mal à comprendre. Cependant, ici aussi les premières étapes concernent la description des fleurs (Figure 1.30).

Pour l'identification d'espèces d'arbre en France métropolitaine uniquement, l'interface web *Key to Nature* (réseau Tela Botanica) propose de répondre à des clefs dichotomiques en fonction de la ressemblance des attributs de l'espèce inconnue à des photos (Figure 1.30).

Un autre logiciel, appelé *IDAO* [3], permet de renseigner les attributs de la plante sans

1. http ://www.tela-botanica.org/eflore-test/bonnierpda
2. http ://id-botanica.com/
3. http ://idao.cirad.fr/



Végétaux ligneux. Arbres à chatons
48

Noisetier, coudrier
Corylus avellana

C Arbrisseau (atteint 6 m). Écorce luisante, brunâtre ou grisâtre, avec verrues (lenticelles). Feuilles alternes (jusqu'à 10 cm) pointues, arrondies à orbiculaires, dentées, un peu velues dessous (pétiole velu-glanduleux). Dioïque. Chatons mâles (jusqu'à 10 cm à maturité), par 1-4 au bout des rameaux ou à l'aisselle des feuilles ; ils fleurissent en février-mars avant les feuilles ; fleurs femelles (6-7 mm) peu apparentes, aspect de bourgeons avec styles rouges, par 1-3 au bout des rameaux courts. Anémogamie. Fructification d'août à octobre. Fruits (noisettes) d'abord blanc verdâtre puis brun-roux à maturité, à coque dure entourée de bractées foliacées. Dispersés par écureuil, geai, sittelle, etc.
H Forêts de feuillus, haies, bosquets. Localisé dans le Midi méditerranéen.
P Atteint 100 ans. Petits chatons mâles visibles dès l'été.

Charme
Carpinus betulus

C Arbre (atteint 25 m). Écorce lisse, grise, tronc cannelé. Feuilles alternes, vert foncé (pétiole : 10-15 mm) dentées, gaufrées dessus, elliptiques (jusqu'à 10 cm), pointues. Monoïque. Les chatons unisexués paraissent en avril-mai avec les feuilles ; les mâles (jusqu'à 7 cm) sont pendants sur des rameaux peu feuillés ; les femelles (3 cm) au bout de rameaux feuillés. Anémogamie. Fruits ailés en grappe (atteignant 14 cm) comprenant 4-10 paires de graines solitaires entourées d'une bractée trilobée, d'abord verts, puis jaunes, asymétriques. Dispersés par le vent et des animaux.
H Forêts de feuillus ; sous les chênes (supporte bien l'ombre). Planté, forme de bonnes haies (charmilles). Rejette de souche.
P Bois blanchâtre, dur, dense, utilisé pour faire des manches, etc ; bon combustible. Croissance rapide, atteint 150 ans.

FIGURE 1.29 – Extrait du livre "Guide de la faune et de la flore" [Eis17].

FIGURE 1.30 – Interface des outils *ID-Botanica* (gauche) et *Key to Nature* (droite).

hiérarchie les uns par rapport aux autres. L'utilisateur peut donc renseigner le type de feuille, puis la couleur de l'écorce, puis la forme de la pointe de la feuille, puis la taille de la plante, et ainsi de suite, jusqu'à ce que le nombre d'espèces végétales correspondantes soit suffisamment réduit. De plus cet outil utilise uniquement des dessins, donc aucun vocabulaire technique comme nous pouvons le voir sur la Figure 1.31.

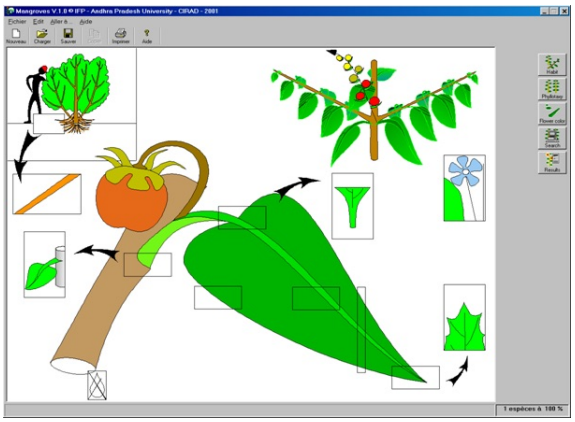


FIGURE 1.31 – Illustration du logiciel IDAO

Clefs sur smartphone

Pour permettre une reconnaissance "sur le terrain" des végétaux, certains de ces outils numériques ont été transposés sur smartphone (applications). Nous pouvons citer *Sauvages de ma rue*[4] ou *IK-Botanic*[5] qui permettent d'identifier des plantes (arbres feuillus, conifères, plantes herbacées ou à spores) selon des critères visuels majoritairement morphologiques. *Quelle est cette fleur ?*[6] permet d'identifier plus de 800 fleurs à partir d'informations sur la couleur, le milieu écologique de la plante (plaine, montagne, aquatique, *etc.*) et du nombre de pétales. Certaines applications sont spécialisées pour la reconnaissance d'arbres, comme *Clés de forêt*[7] de l'Office National des Forêts (ONF) qui propose d'identifier 32 arbres en fonction de leurs feuilles ou de leurs fruits. Le point commun de ses applications est qu'elles fonctionnent sur des critères visuels présentés sous forme de dessins et correspondant à des clefs de détermination botaniques renseignées par l'utilisateur, ce qui est un grand avantage car aucune connaissance botanique n'est requise (Figure 1.32).

Sauf exception de *Quelle est cette fleur ?*, ces applications fonctionnent sans accès à Internet. Un des inconvénients majeurs de ces applications est le peu d'espèces reconnues. Uniquement les arbres les plus communs et qui ont peu de chance d'être confondus avec d'autres espèces sont proposés. Un autre inconvénient est plus un aspect psychologique

4. http ://sauvagesdemarue.mnhn.fr/
5. http ://www.institut-klorane.org/application-IK-BOTANIC
6. www.whatsthatflower.com
7. http ://www.onf.fr/

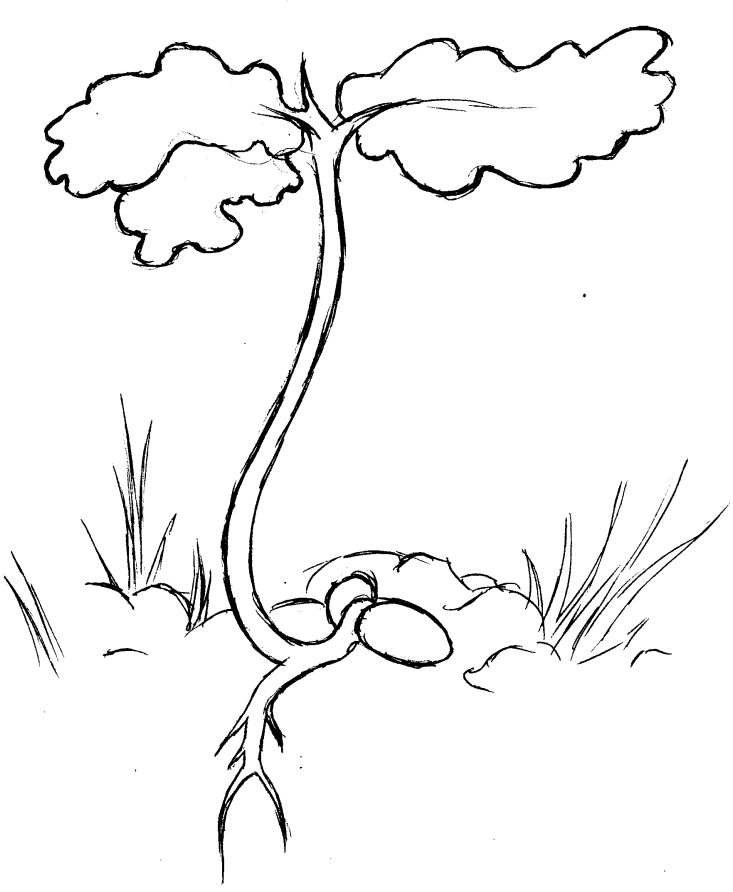
(a) (b) (c) (d)

Figure 1.32 – Captures d'écran des applications pour smartphone *Quelle est cette fleur ?* a), *Sauvages de ma rue* b), *IK-Botanicc*) et *Clés de forêt* d).

de l'utilisateur. En effet, celui-ci peut douter de ses choix, particulièrement sur des clefs où les aspects morphologiques de l'organe sont semblables. Il n'ose plus choisir de peur de se tromper.

Pour pallier ces aspects, des applications pour smartphone ont été développées afin d'identifier automatiquement les espèces végétales. Les applications automatiques répondent à plusieurs problèmes, notamment de pouvoir procéder à une analyse morphologique complexe sans requérir d'expertise de l'utilisateur, de pouvoir prendre en compte différents critères sans être "bloqué" par la progression dans une clé, ou encore de transmettre des informations pertinentes pour l'utilisateur, voire des connaissances pour le néophyte. Ces applications intégrant spécialement des méthodes de traitement d'images vont être détaillées dans le chapitre suivant.

Chapitre 2
État de l'art



L'objectif de ce chapitre est, dans un premier temps, de recenser les différentes applications pour smartphone d'aide à l'identification végétale intégrant des algorithmes de traitement d'images. Par la suite, les différents algorithmes de la littérature traitant de segmentation, d'extraction de caractéristiques et de classification appliqués aux organes végétaux seront détaillés et comparés. L'état de l'art pour l'identification des espèces végétales s'est largement étendu à la suite des Challenges PlantCLEF, qui est un défi visant à évaluer les méthodes et systèmes d'identification en traitement d'images des plantes à très grande échelle, proches des conditions d'un scénario réel telles que la surveillance de la biodiversité. Pour finir, nous établirons les contraintes liées au cadre de cette thèse.

2.1 Systèmes automatiques de reconnaissance de végétaux

Nous avons vu les contraintes liées à l'utilisation des outils manuels pour l'identification des espèces végétales. Aussi, d'autres applications vont plus loin en se basant sur l'analyse automatique ou semi-automatique de photos des organes prises par l'utilisateur pour identifier les végétaux. C'est le cas des applications *LeafSnap*[1], *Pl@ntNet*[2], *GFinger*[3] ou *Folia*[4]. Nous pouvons séparer ces applications en deux catégories selon les méthodes utilisées pour identifier les bonnes espèces végétales.

2.1.1 Applications basées sur l'indexation d'images par le contenu

Les applications les plus performantes actuellement sont celles se basant sur des méthodes de recherche d'image basée sur le contenu (Content-based image retrieval, CBIR), telles que *Pl@ntNet* et *GFinger*. Les méthodes de CBIR consistent à rechercher dans une grande base de données l'image qui correspond au mieux à l'image donnée en entrée à l'algorithme. La recherche s'effectue sur les données visuelles (couleurs, textures, points d'intérêts, *etc.*) plutôt que sur les métadonnées (mots-clefs, étiquette, *etc.*).

L'application *GFinger*

L'application chinoise *GFinger* propose à l'utilisateur une aide pour bien s'occuper de ses plantes (arrosage, entretien) grâce à des fiches techniques et un réseau social de jardiniers amateurs, mais offre aussi un outil automatique de reconnaissance pour identifier les plantes. Celui-ci permet de reconnaître des plantes à partir de photographies des fleurs, des fruits ou des feuilles, mais pas à partir de photos d'écorce (Figure 2.1).

1. http ://leafsnap.com/
2. http ://m.plantnet-project.org/
3. https ://www.lrgarden.com/
4. http ://liris.univ-lyon2.fr/reves/folia

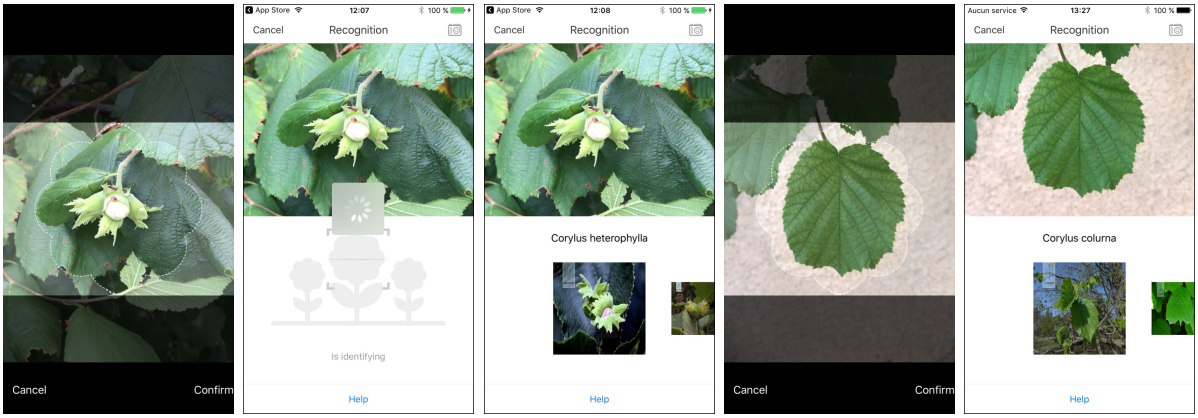


Figure 2.1 – Captures d'écran de l'outil de reconnaissance de plantes de l'application *GFinger*.

L'application *Pl@ntNet*

L'application d'identification d'espèces végétales la plus célèbre en France Métropolitaine est *Pl@ntNet*. Elle peut identifier automatiquement des espèces d'arbres mais aussi des plantes herbacées à partir de photos de feuilles, fruits, fleurs et tige (écorce). L'utilisateur doit juste indiquer à l'application quel type d'organe est présent sur la photo. Dans sa dernière version de 2016, *Pl@ntNet* est capable d'identifier 6220 espèces végétales originaires d'Europe de l'Ouest [Aff+17].

Bien qu'ayant commencé avec des vecteurs de caractéristiques fabriqués "à la main", l'application *Pl@ntNet* exploite depuis 2015 les réseaux de neurones à convolution (Convolutional Neural Networks, CNN) [Cha+15] et a ainsi vu son taux de bonne identification significativement augmenter [Aff+17].

Pour accroître les chances de bonne identification d'une espèce végétale, il est possible de fournir à l'application des photos des différents organes de la plante, c'est-à-dire les feuilles, la tige (le tronc), les fleurs ou les fruits. Si des photos des différentes parties de la plante d'espèce inconnue sont présentées à l'application, les informations extraites sont fusionnées [Cha+15]. Un exemple du processus d'identification via *Pl@ntNet* est présenté Figure 2.2.

Les inconvénients majeurs de ces applications sont avant tout la taille en mémoire de la base de données nécessaire à l'identification, qui interdit une version intégralement embarquée et nécessite donc d'avoir un accès internet pour pouvoir envoyer la photo à traiter sur un serveur qui renverra une réponse à l'utilisateur, ainsi que le temps de calcul qui peut être important.

2.1.2 Applications basées sur le traitement d'images

La seconde catégorie regroupe les applications qui intègrent des algorithmes de traitement d'images permettant d'extraire une signature, ou un vecteur caractéristique, de

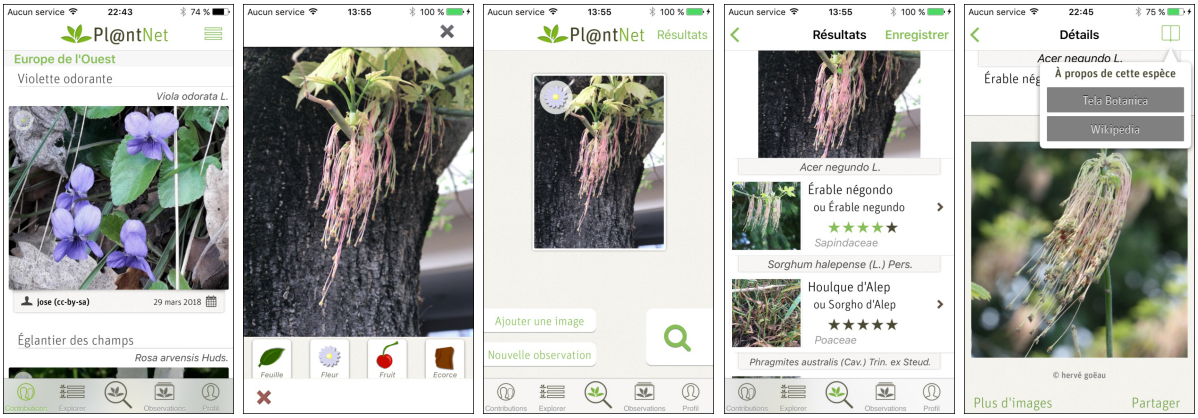


FIGURE 2.2 – Captures d'écran de l'application *Pl@ntNet* lors des différentes étapes d'identification d'une plante.

l'image pour l'affecter à un label (l'espèce correspondante ici). Utilisant cette méthode, nous pouvons citer les applications *LeafSnap* et *Folia*. Ces deux applications identifient des espèces d'arbres et d'arbustes grâce à des photos de leurs feuilles.

L'application *LeafSnap*

LeafSnap[5] est une application gratuite pour smartphone développée en 2011 par l'Université de Columbia, l'Université du Maryland et de l'institut Smithsonian. Elle permet l'identification d'espèces d'arbres présents sur la côte nord-est des État Unis d'Amérique et du Canada (220 espèces pour la dernière version 1.08). Depuis 2014, cet outil de reconnaissance est dupliqué pour traiter les arbres présents au Royaume Uni (*LeafSnap UK*, 156 espèces reconnues dans la dernière version 1.1).

Pour identifier un arbre, *LeafSnap* demande à l'utilisateur de prendre en photo une feuille de l'arbre. Cependant, et c'est l'inconvénient majeur, il est nécessaire que la feuille soit photographiée sur un fond blanc. D'autre part, la photo doit être envoyée sur un serveur qui va faire le traitement (sur le contour), nécessitant donc une connexion Internet. Une fois le traitement effectué, l'application retourne une liste d'espèces probables. L'utilisateur peut alors sélectionner une des espèces proposées pour voir des photos des différents organes (feuilles, fleurs, fruits, écorce) sous différents angles pour vérifier s'il s'agit de la bonne espèce ou non (Figure 2.3).

L'application *Folia*

L'application *Folia* est une application gratuite pour smartphone disponible au téléchargement depuis 2012 permettant l'identification de 95 espèces d'arbres poussant naturellement en France Métropolitaine à partir d'une photo de leurs feuilles.

Contrairement à *LeafSnap*, il n'est pas nécessaire de prendre la feuille sur un fond

5. http ://leafsnap.com

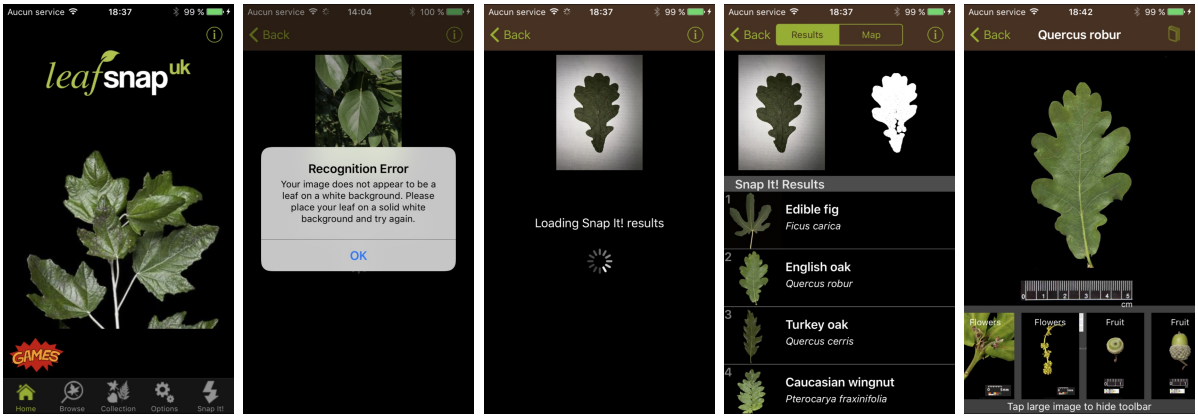


Figure 2.3 – Captures d'écran de l'application *LeafSnap UK* lors des différentes étapes d'identification d'un arbre.

blanc, et tous les traitements (segmentation, extraction de caractéristiques, et classification) sont effectués sur le téléphone. Il n'est donc pas nécessaire d'avoir un accès à Internet, ce qui peut arriver dans un scénario dans lequel l'utilisateur souhaite identifier des arbres rencontrés au cours d'une promenade à la campagne ou en montagne. Contrairement aux applications précédemment présentées, *Folia* demande deux interactions de la part de l'utilisateur. En effet, celui-ci doit indiquer s'il s'agit d'une feuille simple ou d'une feuille composée, et dessiner une trace à l'intérieur de la feuille, ou des folioles dans le cas d'une feuille composée, pour désigner les différentes couleurs du limbe à l'algorithme. Une fois ces étapes franchies, et au bout des différentes phases de traitement d'images, l'application propose au botaniste amateur une liste de 10 espèces qui ont le plus de chances de correspondre à l'arbre inconnu (Figure 2.4). En sélectionnant une espèce, l'utilisateur a accès à une fiche espèce avec une description et des illustrations. S'il pense avoir trouvé l'espèce correspondante, l'utilisateur peut valider celle-ci et conserver l'image dans son herbier numérique.

Figure 2.4 – Captures d'écran de l'application *Folia* présentant les interfaces d'interaction utilisateur.

Dans la suite de ce chapitre, nous allons voir les différents algorithmes utilisés par ces applications, ainsi que les méthodes de traitement d'images de la littérature appliquées à la reconnaissance de végétaux.

2.2 Détection et segmentation des différents organes de l'arbre

En traitement d'images, la segmentation vise à regrouper des pixels en régions homogènes, pour partitionner l'image. La plupart du temps, nous cherchons à diviser l'image en deux classes : une région regroupant les pixels de l'objet d'intérêt, et une région regroupant les pixels de l'arrière-plan, afin d'effectuer une binarisation. Ainsi, il devient plus facile d'extraire des informations uniquement sur l'objet d'intérêt et non sur l'arrière-plan. Pour ce qui est de l'identification des espèces d'arbres et arbustes, les objets d'intérêt que nous pouvons chercher à segmenter sont la feuille, le fruit, l'inflorescence ou l'écorce d'un arrière-plan plus ou moins complexe. Dans la suite, nous allons voir un ensemble de méthodes de segmentation de la littérature qui sont appliquées aux organes végétaux.

2.2.1 Segmentation de l'utilisateur

Segmentation manuelle

Il existe des méthodes de segmentation simples et générales qui peuvent s'appliquer à tous les types d'organes de la plante. La première est la segmentation manuelle par l'utilisateur. L'algorithme demande alors à l'utilisateur de séparer lui-même l'objet d'intérêt de l'arrière-plan. C'est la méthode utilisée par Casanova *et al* [Cas+12] pour ensuite détecter le contour de feuilles photographiées en milieu naturel.

Cette méthode demande une grande implication de l'utilisateur et de définir manuellement les coordonnées d'un grand nombre de points pour avoir une segmentation de bonne qualité, et n'est donc absolument pas automatique.

De manière plus simple, pour une application de reconnaissance d'écorce d'arbre, la méthode mise au point par Huang *et al.* [Hua+06a] demande à l'utilisateur de tracer un rectangle définissant une région d'intérêt (ROI - Region Of Interest) sur une photo de tronc. À cause de sa forme cylindrique, les côtés peuvent être flous. En définissant une ROI, uniquement la texture nette est conservée.

Segmentation par Ligne de Partage des Eaux

Demander à un utilisateur de segmenter entièrement une feuille ou une fleur peut être assez contraignant. Mais d'autre part, il peut être difficile de créer un algorithme de segmentation sans initialisation. L'algorithme de la Ligne de Partage des Eaux (LPE) [BM92]

suppose qu'une image en niveau de gris peut être considérée comme un relief topologique, où les niveaux de gris correspondent aux hauteurs. L'objectif est alors de trouver la ligne de partage des eaux, c'est-à-dire l'endroit où se rejoignent deux bassins versants. À partir de points sources, les bassins sont inondés. Lorsque deux bassins de sources différentes se rejoignent, une digue est créée. Lorsque cette méthode est utilisée pour segmenter un objet d'intérêt, il est demandé à l'utilisateur de fixer les sources (marqueurs). Un marqueur interne est placé dans l'objet à segmenter et un ou plusieurs marqueurs externes dans l'arrière-plan. C'est cette technique qui est utilisée par Le *et al.* [Le+15] et Nguyen *et al.* [Ngu+17] pour segmenter une feuille unique présente sur un fond végétal, ainsi que par Lee et Hong [LH17] pour la segmentation de fleurs. La Figure 2.5 illustre l'algorithme LPE.

Figure 2.5 – Interface pour la segmentation d'une feuille, avec un marqueur interne et des marqueurs externes, les bassins versants générés, et la feuille segmentée. Extrait de Le *et al.* [Le+15]

Segmentation interactive

Une segmentation interactive est une segmentation qui permet des interactions entre l'algorithme et l'utilisateur.

Ainsi, la méthode de segmentation de fleur de Hsu *et al.* [HLC11] demande à l'utilisateur de tracer une boîte englobante autour de la fleur qu'il souhaite segmenter, pour initialiser l'algorithme de segmentation qui va venir trouver les contours de celle-ci grâce à un opérateur de Sobel [Sob14]. Quatre points sont créés sur le milieu de chaque côté de la boîte englobante qui vont venir se déplacer sur la ligne reliant deux points opposés, jusqu'à atteindre le contour de la fleur. Si les points ne se situent pas sur le contour de la fleur, l'utilisateur peut ajuster leur position (Figure 2.6).

L'ensemble du contour de la fleur est ensuite calculé en reliant les points en minimisant une fonction de coût. Dans cet article, les quatre points sont placés automatiquement. Mais

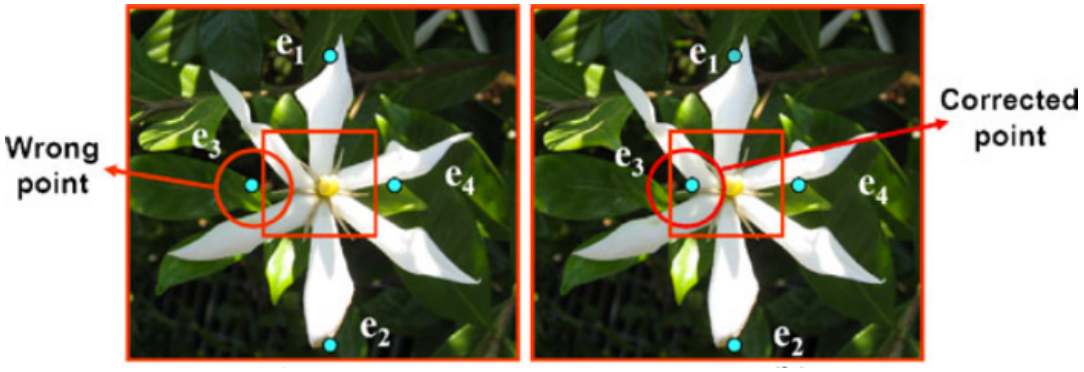


FIGURE 2.6 – Boîte englobante (cadre extérieur) autour de la fleur à segmenter et détection, puis correction, des points de départ de la création du contour, extrait de Hsu *et al.* [HLC11]

avec la méthode des ciseaux intelligents [MB95], l'utilisateur peut placer des points (appelés "ancres") le long du contour de l'objet d'intérêt qui seront ensuite reliés en minimisant une fonction de coût se basant sur les contours (mesure du gradient). Si le contour obtenu avec les ciseaux intelligents ne suit pas bien les bords de l'objet d'intérêt, l'utilisateur peut rajouter des points sur la frontière de l'objet pour que le contour soit recalculé entre les points adjacents. Une illustration de l'utilisation des ciseaux intelligents est présentée sur la Figure 2.7.

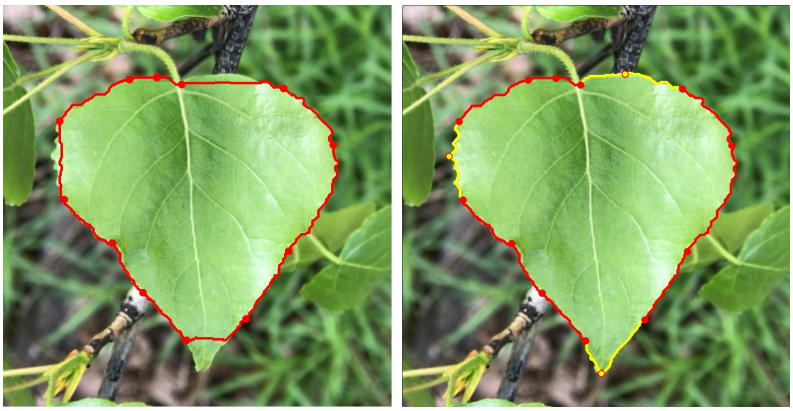


FIGURE 2.7 – Utilisation de l'algorithme des ciseaux intelligents [MB95] pour la segmentation d'une feuille, avec ajout de trois ancres lorsque le contour extrait ne suit pas bien le bord de la feuille.

Un autre outil de segmentation interactive est l'algorithme GrabCut [RKB04]. GrabCut est une méthode de segmentation d'image basée sur des coupures de graphes. L'utilisateur délimite une boîte englobante autour de l'objet à segmenter. L'algorithme estime la distribution des couleurs de l'objet d'intérêt et de l'arrière-plan à l'aide d'un modèle de mélange de gaussiennes. Ceci est utilisé pour construire un champ aléatoire de Markov sur les étiquettes des pixels, avec une fonction d'énergie qui préfère les régions connectées ayant la même étiquette. Les estimations peuvent être corrigées par l'utilisateur en indiquant les régions mal classées, comme représenté sur la Figure 2.8. Cet outil de seg-

mentation a notamment été utilisé par Arora *et al.* [Aro+12] pour segmenter des feuilles sur fond complexe naturel.

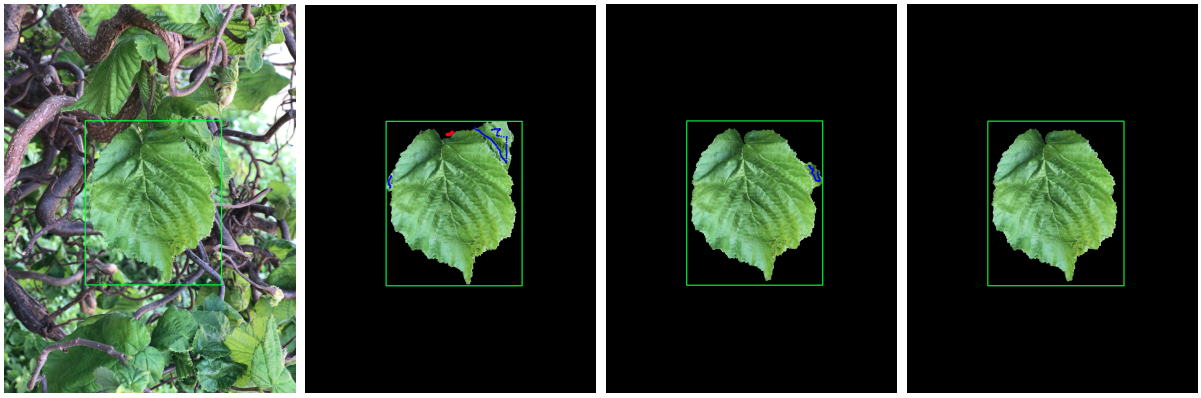


Figure 2.8 – Les différentes étapes de la segmentation d'une feuille avec GrabCut. Une première segmentation est calculée à partir de la ROI tracée par l'utilisateur (rectangle vert). L'utilisateur peut ensuite rajouter des parties manquantes (points rouges) et indiquer des parties n'appartenant pas à l'objet d'intérêt (points bleus).

Pour finir avec les segmentations interactives, nous pouvons citer le travail de Zou et Nagy [ZN04], qui proposent une interface pour segmenter des fleurs prises de face. L'équation légèrement modifiée d'une rosace (*rhodonea*) est appliquée pour modéliser la fleur. L'utilisateur peut régler les paramètres de cette équation, correspondant au centre, au rayon externe et interne, au nombre de pétales et à l'orientation de la fleur, afin que la courbe s'ajuste au mieux au contour de la fleur (voir Figure 2.9).

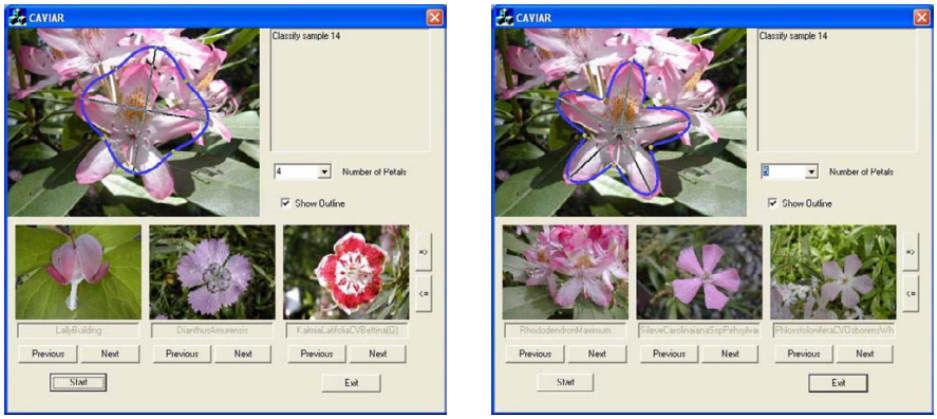


Figure 2.9 – Illustration de l'ajustement des paramètres de la courbe *rhodonea* pour la segmentation d'une fleur, issue de [ZN04].

La segmentation précise d'un objet d'intérêt à la main peut être un processus assez long et pénible pour un utilisateur. Nous avons donc vu, dans le cadre de la segmentation d'éléments botaniques, des algorithmes facilitant le travail de segmentation de l'utilisateur. L'avantage de l'interactivité entre l'algorithme et l'utilisateur est d'obtenir une segmentation correcte en cas d'erreur de l'algorithme.

2.2.2 Segmentation automatique

Enfin, il existe également des méthodes de segmentation automatiques ne demandant aucunes interactions avec l'utilisateur. Nous allons présenter dans les prochains paragraphes des algorithmes automatiques de segmentation qui ont été utilisés dans la littérature pour l'extraction des organes végétaux.

Seuillage

Une méthode assez simple et classique, pas uniquement applicable sur les organes végétaux, est le seuillage d'image. Il s'agit d'une méthode de binarisation de l'image, où chaque pixel de l'image est remplacé par un pixel noir si la valeur de celui-ci est inférieure à un seuil, et remplacé par un pixel blanc si sa valeur est supérieure au seuil. Le seuillage peut se faire en fonction du niveau de gris [Wu+07], sur la couleur (notamment en imagerie hyperspectrale ou multispectrale, imagerie thermique, lumière structurée) [LLH11], ou sur la texture [ZTK05]. La valeur du seuil peut être fixée empiriquement, notamment lorsque la prise de vue est déterminée. Par exemple, la base de données de feuilles de Wu *et al.* [Wu+07] est constituée de feuilles scannées. Les images sont premièrement converties en niveau de gris, puis seuillées avec une valeur fixe de niveau de gris.

Toutefois, la valeur du seuil peut être calculée pour chaque image, afin d'être un peu plus automatique et optimal en fonction de l'image. C'est très souvent la méthode de Otsu [Ots79] qui est utilisée. Cet algorithme suppose que l'image en niveau de gris est composée de deux classes. Le principe est de minimiser la variance intra-classe des deux classes séparées par un seuil. Cette méthode a été utilisée pour segmenter des feuilles à partir d'images en niveau de gris [Mzo+16], des fleurs en se basant sur les histogrammes de couleur [NZ12], ou des fruits [Jid+16]. Cependant, cette méthode est vite limitée quand il n'y a pas de séparation large entre les pics des deux classes sur l'histogramme. Ceci peut arriver lorsque l'objet d'intérêt est trop petit par rapport au reste de l'image, ou s'il y a peu de contraste entre l'objet d'intérêt et l'arrière-plan.

Une autre méthode de segmentation régulièrement utilisée est la méthode de clustering dite des k-means. Il s'agit d'un outil utilisé pour partitionner l'image en k clusters, fonctionnant de la manière suivante : chacun des pixels de l'image est représenté dans un espace vectoriel contenant k centroïdes. De manière itérative, chaque pixel est assigné au centroïde le plus proche, puis les positions des k centroïdes sont recalculées en fonction des vecteurs des pixels affectés à chaque cluster, jusqu'à convergence. Une image peut être binarisée si l'on considère uniquement deux clusters (un correspondant à l'objet d'intérêt et un autre pour l'arrière-plan). Cette méthode a été utilisée pour segmenter des fleurs ou des feuilles photographiées sur fond noir [SK00], ou des fruits photographiés sur une table de couleur claire [DJ15].

La segmentation de l'organe d'intérêt de la plante se complexifie quand celui-ci ne se

détache pas de l'arrière-plan, par exemple s'il y a plusieurs feuilles présentes sur l'image, ou si la fleur ou le fruit est de même couleur que l'arrière-plan, comme c'est le cas par exemple pour l'espèce d'arbre *Acer campestre* (Figure 2.10).

Figure 2.10 – Les différents organes de *Acer campestre* sont parfois difficiles à isoler de l'arrière-plan.

Segmentation via la profondeur

L'objet d'intérêt étant le plus souvent au premier plan, l'information de profondeur peut être utilisée pour segmenter les différents organes végétaux. Une méthode coûteuse mais régulièrement utilisée en agriculture est la stéréo-vision. En comparant l'information d'une scène à partir de deux points de vue, l'information 3D peut être extraite en examinant les positions relatives des objets dans les deux images. Cette technique est notamment utilisée pour le développement de robots de ramassage de fruits automatiques [PG01 ; XJY14].

Un système plus abordable est la camera RGB-D, permettant d'obtenir des images intégrant la profondeur, dont le plus célèbre est la *Kinect* de *Microsoft*. L'information de profondeur est obtenue par la projection d'un ensemble de points en infrarouge dans la scène puis par le calcul des déplacements de ces points sur les surfaces à cartographier. Ce genre d'outil a notamment été utilisé pour segmenter des feuilles [Xia+15].

Cependant, le problème de l'utilisation de la 3D est la nécessité d'avoir un capteur de profondeur qui demande un calibrage de l'appareil. Il existe des algorithmes de segmentation fonctionnant avec des images prises par un appareil photo classique [TKC09]. Pour cela, il faut plusieurs images de l'organe pris sous des points de vue légèrement décalés, pour pouvoir évaluer le flux optique et estimer les paramètres de la caméra pour en déduire la profondeur (Figure 2.11). Les images doivent être légèrement différentes pour que la correspondance entre les points d'intérêt des images se fasse correctement.Toutefois, si les photos en entrée sont trop décalées, la correspondance entre les points 2D et 3D ne s'effectue pas correctement et des points sont mal segmentés.

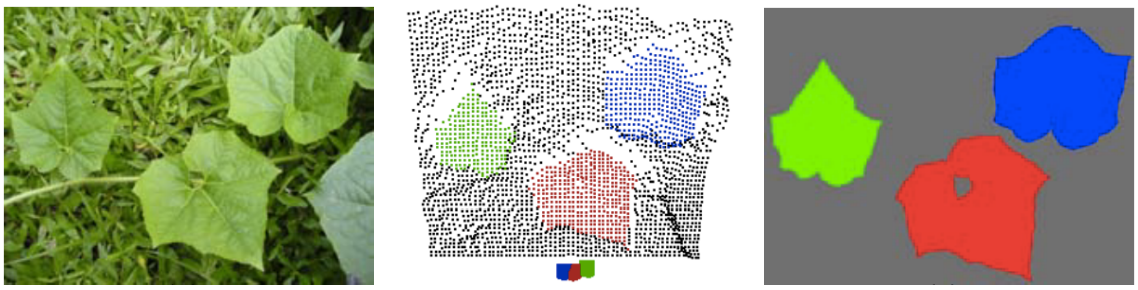


FIGURE 2.11 – Illustration des résultats de l'algorithme de segmentation 3D issue de [TKC09] avec une des images sources, les points 3D appairés et les feuilles segmentées.

Segmentation avec connaissance *a priori* de forme

Au-delà de ces méthodes générales de segmentation, des algorithmes ont été créés spécialement pour s'adapter aux types d'organes de la plante. Cela requiert une connaissance *a priori* du contenu de l'image, c'est-à-dire de savoir s'il s'agit d'une image de feuille, de fleur, de fruit ou d'écorce. Cette connaissance permet de mettre en place des algorithmes de segmentation spécifiques.

Il est ainsi possible de se servir de modèles de forme de feuilles pour guider la segmentation. Manh *et al.* [Man+01] ont décidé d'utiliser un modèle déformable pour segmenter les feuilles, même superposées, d'une seule espèce de plante (*Setaria viridis*). L'utilisation d'un modèle de forme a été reprise par Cerutti *et al.* [Cer+11] pour s'adapter à toutes les formes de feuilles simples ou lobées. En première étape, une carte de dissimilarité de couleur dans l'espace L*a*b* est calculée pour faire ressortir la feuille à segmenter. Dans l'article [Cer+11], la feuille est considérée comme étant au centre de l'image et correctement orientée (apex vers le haut). L'échantillon de couleur est donc pris au centre de l'image, mais il peut aussi être récupéré à partir d'une trace dessinée par l'utilisateur. Ensuite, un modèle polygonal de forme déformable est appliqué sur cette carte de dissimilarité de couleur pour correspondre à la forme de la feuille (Figure 2.12). Ceci permet de contraindre la segmentation à une forme de feuille, pour ainsi isoler la feuille d'intérêt des autres feuilles appartenant à l'arrière-plan.

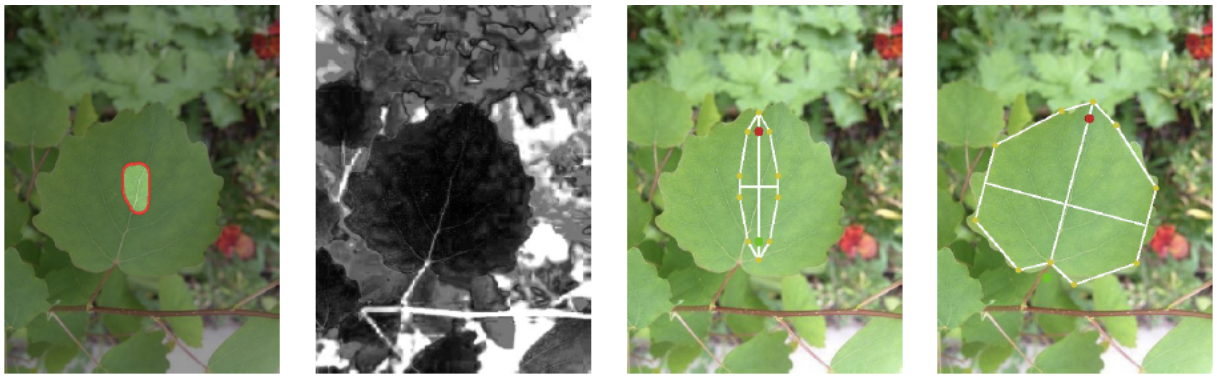


FIGURE 2.12 – Application d'un modèle paramétrique déformable de feuille pour la segmentation, issue de [Cer13b]

Pour finir, l'auteur va utiliser le modèle polygonal pour contraindre une segmentation par contours actifs pour ainsi récupérer le contour et la marge de la feuille. Dans [Gra+15], les auteurs ont pu montrer que cette segmentation surpasse les autres algorithmes de segmentation de la littérature dans ce contexte. C'est cet algorithme qui est implémenté dans l'application *Folia* pour segmenter les feuilles en environnement complexe.

De même que pour la feuille, certains chercheurs utilisent un modèle de forme de fleur pour aider à la segmentation. La publication de Nilsback et Zisserman [NZ07] propose une méthode en deux étapes regroupant les aspects de couleur et de forme. La première consiste à obtenir une segmentation initiale de la fleur en utilisant les distributions de couleur de l'arrière-plan et de l'objet d'intérêt. Ces distributions de couleurs ont été apprises en labellisant les pixels des images d'apprentissage en tant qu'arrière-plan ou objet d'intérêt. La seconde étape consiste à appliquer un modèle de forme de fleur générique sur la segmentation initiale pour détecter les pétales. Les couleurs des pétales sont ensuite utilisées pour ré-ajuster le modèle de distribution des couleurs de la fleur (celui de l'arrière-plan reste inchangé) et la segmentation est de nouveau évaluée. La seconde étape est répétée jusqu'à convergence. Une vue d'ensemble de ces étapes de segmentation est illustrée sur la Figure 2.13.

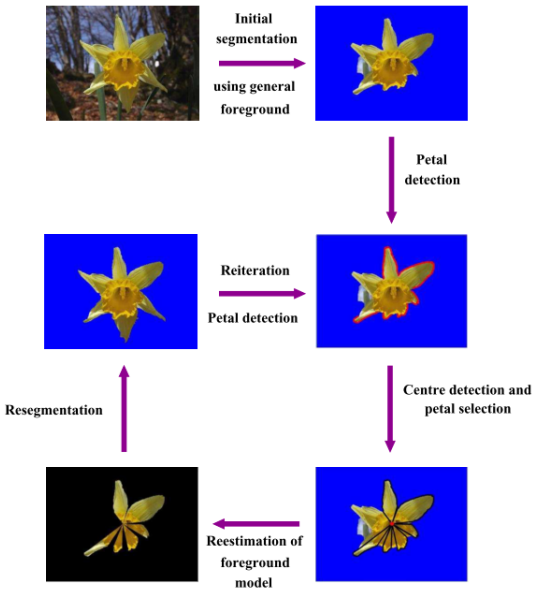


Figure 2.13 – Résumé de l'algorithme de segmentation de Nilsback et Zisserman, issu de [NZ07]

Cependant, cette méthode de segmentation ne fonctionne pas avec des fleurs de couleur verte, car souvent l'arrière-plan est aussi vert.

Les systèmes de segmentation de fruits sont surtout développés pour la détection de fruits sur les arbres, pour des machines de ramassage de fruits. Ces machines sont spécialisées pour la collecte de fruits cultivés pour l'homme, c'est-à-dire des pommes, oranges, pêches, litchis, *etc* [PJJ12], soit des fruits généralement sphériques. Cet aspect a

donc été pris en compte pour améliorer les étapes de segmentation. Patel *et al.* [PJJ12] cherchent dans un premier temps à appliquer un opérateur de Sobel pour détecter les bords du fruit, avant de leur ajuster un cercle (Figure 2.14). De manière semblable, c'est la transformée de Hough qui est utilisée pour détecter les fruits ronds dans [PG01].

FIGURE 2.14 – Illustration de l'utilisation de l'aspect circulaire des fruits pour la segmentation, issue de [PJJ12]

Mettre en place un algorithme de segmentation performant sur tout type d'organe végétal est un vrai challenge. La segmentation par l'utilisateur donne les meilleurs résultats pour peu que celui-ci soit précis. Cependant, cette méthode est chronophage et pénible, en particulier pour des objets de forme complexe. L'introduction de connaissance *a priori* de l'objet à segmenter, via l'utilisation de la profondeur ou des modèles de forme, permet de faciliter la résolution de ce problème complexe de segmentation. Toutefois, un modèle déformable de forme ne peut s'appliquer à l'ensemble des organes végétaux. Sur des images de différents organes de plante photographiés en milieu naturel, donc avec un arrière-plan complexe ne permettant pas un seuillage aisé de l'objet d'intérêt, Joly *et al.* [Bak+13] extraient des caractéristiques locales dans l'image toute entière grâce à un détecteur de points de Harris prenant en compte la couleur. Les points sont extraits sur quatre dimensions pour prendre en compte les différentes échelles et les problèmes de flou. Pour éviter d'extraire un trop grand nombre de points appartenant à l'arrière-plan, un masque parallélépipédique a été appliqué à l'image d'entrée et plus d'importance a été accordée aux points proches du centre de l'image. Après cette étape, les points sont localisés sur l'organe, mais il n'y a pas de réelle segmentation (les caractéristiques locales extraites autour des points faisant office de descripteurs.).

2.3 Extraction des descripteurs

La segmentation d'un objet d'intérêt permet par la suite d'extraire des descripteurs (vecteurs caractéristiques) représentatifs exclusivement de cet objet, que ce soit des informations sur le contour de la forme ou des éléments caractéristiques à l'intérieur de celui-ci.

2.3.1 Descripteurs de forme

Nous avons vu dans le chapitre précédent que la forme de l'organe végétal, que ce soit la feuille, la fleur ou le fruit, est généralement décrite car elle est extrêmement utile pour l'identification de l'espèce. C'est donc naturellement que des descripteurs caractérisant la forme sont utilisés en traitement d'images appliqué à la reconnaissance de végétaux.

Il est aisé, à partir d'une image binaire (issue d'une étape de segmentation), d'extraire le contour de l'objet d'intérêt. Le contour de l'objet est la séquence de pixels qui se trouvent sur son bord.

Il existe des paramètres génériques pour décrire les formes qui sont aussi utilisés pour caractériser les objets, tels que le diamètre, le périmètre, l'aire, la rectangularité *etc*, par exemple pour caractériser les feuilles [Wu+07 ; BKK13], ou des fruits [Zha+14]. La rotondité de la fleur est calculée en utilisant le rapport entre l'aire et le périmètre au carré (paramètre appelé circularité) [SK03]. Ces paramètres ont l'avantage d'être rapides à calculer. Ces descripteurs morphologiques sont efficaces pour différencier les formes qui sont très différentes. Ils sont donc souvent utilisés pour réduire l'espace de recherche. Ils sont invariants à la translation et à la rotation.Mais ils sont trop "globaux" pour être discriminants dans notre cas, à cause de la similarité inter-classe et à la grande variabilité intra-classe.

Des descripteurs de courbure sont aussi employés pour caractériser la forme générale de la feuille et son aspect sinueux, soit la concavité et la convexité des contours. Le descripteur CSS (Curvature Scale Space), qui permet de décrire les convexités (courbures positives) et les concavités (courbures négatives) a été implémenté pour la classification des feuilles [MA04]. Pour chaque point du contour, la courbure est calculée à des échelles croissantes (grâce à un lissage gaussien). Cette méthode a été implémentée pour la classification des feuilles [MA04], et particulièrement dans l'application *LeafSnap* présentée en début de chapitre.

Pour décrire la forme des pétales, Saitoh *et al.* [SK03] utilisent la distance entre le centre de la fleur et le contour, comme illustré sur la Figure 2.15.

À partir du descripteur de contour ainsi construit, la variable l représente la largeur du pétale, le nombre de maxima correspond au nombre de pétales (moins il y a de pétales, plus ceux-ci sont séparés et les périodes faciles à distinguer). Cette méthode se rapproche du descripteur CCD (Centroid Contour Distance) qui calcule une séquence de distances

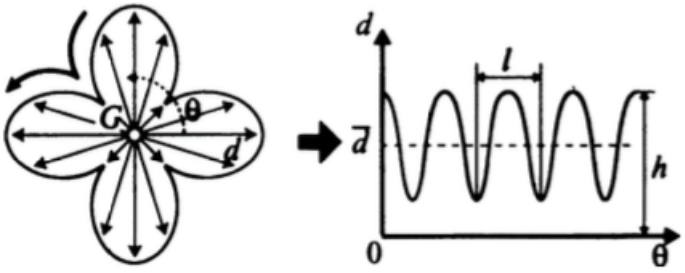


FIGURE 2.15 – Analyse de la forme de la fleur selon Saitoh *et al.*. Illustration issue de [SK03].

entre le barycentre de la forme et les points du contour externe, à différentes échelles, ce qui donne une fonction 1D, qui a été utilisée par Hong *et al* [Hon+03] pour décrire le contour des fleurs, ou par Wang *et al.* [Wan+00] pour décrire la marge des feuilles. Une autre méthode consiste simplement à calculer le nombre de dents, qui sont détectées via les maxima de courbure [Aro+12] (Figure 2.16).

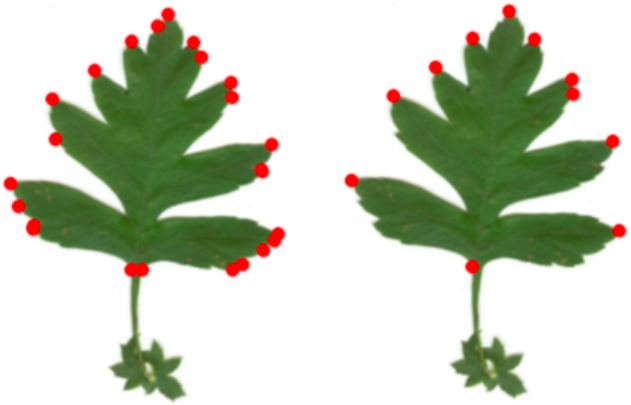


FIGURE 2.16 – Détection des dents sur le contour de la feuille avec deux valeurs de seuil différentes pour les maxima de courbure, issue de [Aro+12]

Particulièrement pour les feuilles, des paramètres de forme spécifiques ont été développés pour analyser la base et l'apex (la pointe) de la feuille, basés sur les courbes de Bézier [Cer+13b].

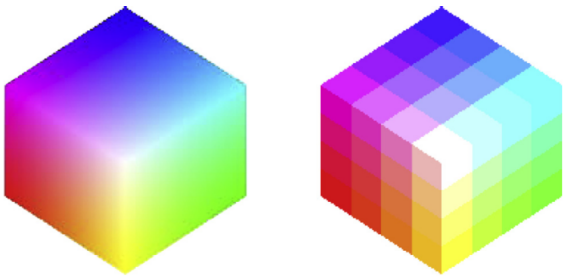
Le modèle paramétrique de forme utilisé pour la segmentation, que nous avons vu dans la partie 2.2.2 peut aussi être utilisé pour la caractérisation des feuilles [Cer+11] et des fleurs [NZ07]. Ainsi le nombre de lobes de la feuille ou le nombre de pétales de la fleur peut permettre la réduction du nombre d'espèces probables.

2.3.2 Descripteurs de couleur et de texture

Si la couleur de la feuille est utilisée lors des étapes de segmentation, celle-ci est très peu employée pour la caractérisation, car la couleur n'est pas très caractéristique d'une

espèce. En effet, la couleur de la feuille est donnée par la présence de chlorophylle dans celle-ci, et dépend de l'ensoleillement, de la région géographique, et surtout de la saison, vu que les feuilles changent de couleur à l'automne. Elle est cependant très utilisée pour les autres organes végétaux.

Pour caractériser la couleur des fleurs, ce sont principalement les histogrammes de couleurs qui sont utilisés. Lors du Challenge PlantCLEF de 2013, l'équipe de l'INRIA [Bak+13] a employé un histogramme RGB pondéré de 27 bins et un histogramme HSV de 30 bins. La plupart du temps, c'est l'espace colorimétrique HSV qui est utilisé pour la description de la couleur chez les fleurs [NZ08] car dans celui-ci la teinte n'est codée que sur un seul canal, rendant la description de la teinte moins sensible aux variations de luminosité et donc aux conditions de prise de vue. Les histogrammes de couleurs sont généralement de grandes dimensions, alors qu'il n'est parfois pas nécessaire d'avoir toutes les teintes de couleur pour faire la différence entre par exemple une pomme et une banane. Zhang *et al.* [Zha+14] proposent de discrétiser l'espace de couleur RGB avec 4 bins par canal (Figure 2.17).



(a) 256×256×256 RGB Space (b) 4×4×4 RGB Space

Figure 2.17 – Illustration de la discrétisation de l'espace colorimétrique RGB avec 4 bins par canal, issue de [Zha+14].

Une autre méthode pour caractériser la couleur des fleurs ou des fruits d'une espèce végétale peut être de créer un vocabulaire de couleurs (un sac de mots). Dans [Cer+13a], le vocabulaire est appris en clusterisant les valeurs de couleurs de l'espace L*a*b* des différentes fleurs ou différents fruits segmentés (Figure 2.18).

Comme nous l'avons vu dans la partie 2.2 précédente, c'est surtout pour les fruits consommés par l'homme que des algorithmes de détection, et ensuite de reconnaissance, existent. Les algorithmes se concentrent donc sur les couleurs des fruits pour les identifier [Ari+10 ; Bak+13]. La texture des fruits permet elle aussi leur identification. La texture de la fleur est étudiée, avec par exemple des outils d'analyse de texture tels que les matrices de co-occurences en niveau de gris ou les filtres de Gabor [GSM10]. Nilsback et Zisserman ont mis en place un vocabulaire de textures à partir d'une banque de filtres pour la reconnaissance de fleurs [NZ06].

La reconnaissance des espèces d'arbres *via* l'écorce est un problème relativement récent en traitement d'images. Le nombre de publications est assez restreint contrairement à celui

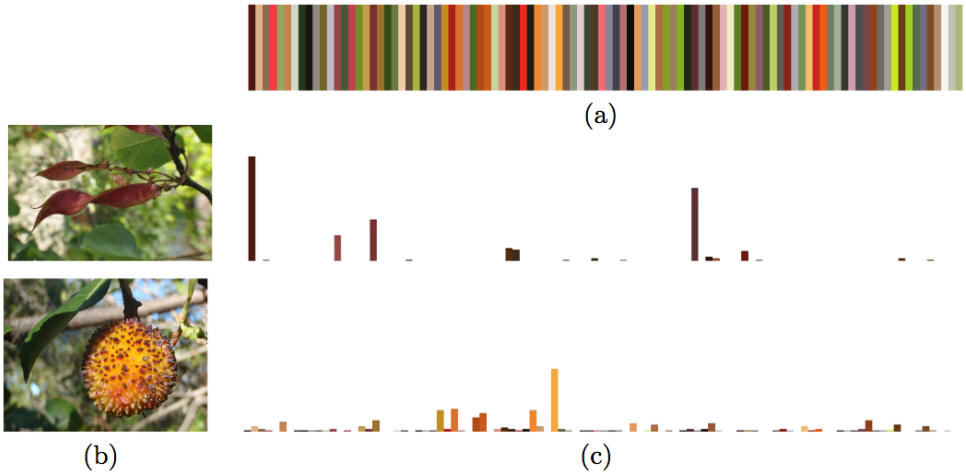


FIGURE 2.18 – Vocabulaire de couleurs appris pour la reconnaissance de fruits ou de fleurs, issue de [Cer+13a].

concernant les feuilles et les fleurs.

La reconnaissance d'écorce peut être vue comme un problème d'analyse de textures, puisqu'il s'agit de caractériser la structure du tronc, et non sa forme. La publication de Wan *et al.* [Wan+04] reprend et compare différents outils d'analyse statistique de textures utilisés pour la reconnaissance d'écorce d'arbres. Les méthodes généralement proposées se basent sur l'utilisation de matrices de co-occurence, des méthodes d'auto corrélation, l'analyse de l'histogramme des niveaux de gris, ou des matrices de longueur de plage (run-length method RLM). L'emploi de matrices de co-occurence sur des images d'écorces en niveaux de gris est aussi présent dans [Hua+06b] où Huang *et al.* améliorent la reconnaissance par rapport aux méthodes précédentes en les combinant avec des dimensions fractales caractérisant l'auto-similitude de la texture de l'écorce à différentes échelles.

Les méthodes spectrales telles que les filtres de Gabor sont aussi appliquées, comme introduit dans l'article de Huang *et al.* [Hua+06a], qui démontrent que seulement quatre orientations d'onde et 6 échelles sont suffisantes pour identifier des espèces d'arbre via leurs écorces. Les filtres de Gabor ainsi que les matrices de co-occurrences ont aussi été utilisés pour l'identification des espèces d'arbres à partir de feuilles [CPB15]. C'est spécialement les nervures qui donnent l'aspect texturé des feuilles. Leur orientation, leur courbure et leur fréquence peuvent être caractérisées par des ondelettes de Gabor [KYU10].

Une autre méthode d'analyse de textures couramment utilisée pour la reconnaissance d'écorce d'arbres s'appuie sur les LBP (Local Binary Patterns). Boudra [BYB15] emploie une multitude de LBP pour classifier les écorces. L'avantage des LBP est l'invariance à la luminance.

La plupart des méthodes statistiques citées ci-dessus ont été appliquées sur des images d'écorce en niveaux de gris. Pour ne pas perdre les informations fournies par la couleur, Wan *et al.* [Wan+04] appliquent les méthodes testées dans son article sur chacun des canaux de l'espace RGB. La prise en compte de la couleur augmente le taux d'identi-

fication des espèces d'arbre. Différents espaces colorimétriques sont utilisés, notamment l'espace HSV dans [Bak+13] où les auteurs manient divers types d'histogrammes pour caractériser l'information de couleur. Il a aussi été montré que la combinaison des informations de couleurs issues de l'histogramme HSV et des informations de texture améliore la reconnaissance des types de fruits [Ari+10].

2.3.3 Descripteurs locaux

Contrairement aux descripteurs globaux tels que la texture ou la couleur qui décrivent l'ensemble de la région d'intérêt, les descripteurs locaux vont caractériser uniquement des petites zones ou des points d'intérêt saillants. Parmi les plus célèbres, nous pouvons citer le descripteur SIFT (Scale Invariant Feature Transform), le descripteur SURF (Speeded-up Robust Features), caractérisant de manière invariante à l'échelle et à la rotation des points d'intérêt (appelés keypoints).

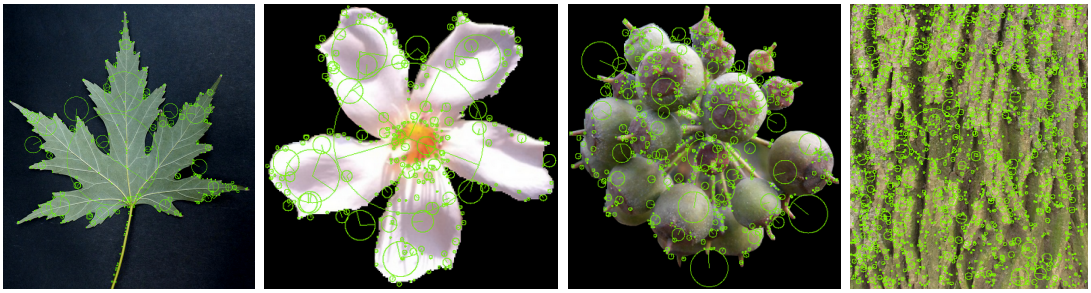


Figure 2.19 – Visualisation des points SIFT sur différents organes végétaux.

Pour Nilsback et Zisserman [NZ08], le contour de la fleur segmentée peut être défini en calculant le descripteur SIFT tout le long du pourtour avec un certain pas. L'intérieur de la fleur est aussi décrit avec le descripteur SIFT (en niveau de gris) calculé à toutes les intersections d'un quadrillage superposant la fleur. Un vocabulaire est enfin créé en regroupant par un algorithme de clustering les vecteurs SIFT issus du contour de la fleur, et un autre vocabulaire à partir des vecteurs issus de l'intérieur. Ce descripteur a également été utilisé pour caractériser les textures d'écorces et les feuilles/aiguilles [WSG11].

L'état de l'art sur la reconnaissance de fruit n'est pas très important car la plupart des équipes de recherche utilisent des descripteurs de points d'intérêts [Nak13][Bak+13], c'est-à-dire SIFT ou SURF, en les regroupant parfois en histogramme de sac-de-mots [ILP13] [Cer+13a].

2.4 Classification

Une fois que les vecteurs descripteurs ont été calculés sur les images, il faut assigner ce vecteur (et ainsi l'image) à une classe. Cette classe peut représenter une famille, un genre, une espèce ou un groupe d'espèces.

Pour mimer les clefs de détermination botaniques, le système de classification peut être un arbre de décision. Il peut être représenté sous forme graphique par un arbre représentant un ensemble de choix (Figure 2.20).

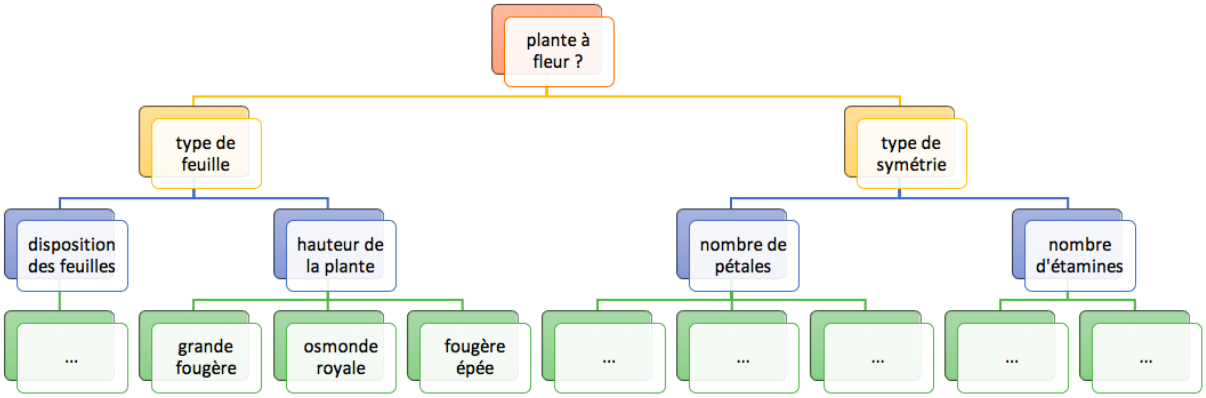


FIGURE 2.20 – Exemple d'arbre de décision.

Avec un apprentissage supervisé, il est possible de calculer automatiquement les variables discriminantes du vecteur caractéristique. Un arbre de détermination serait optimal pour apprendre à un utilisateur novice en botanique à reconnaître les espèces avec des clefs botaniques. Cependant, comme nous l'avons vu dans le chapitre précédent, ces clefs dichotomiques existantes se réfèrent à des organes qui ne sont pas toujours présents (fleurs ou fruits) et à des descriptions de parties invisibles sans dissection. Nous pourrions créer un arbre de décision construit à partir de nos données, mais celui-ci ne respecterait pas la systématique et ne serait pas très précis. De plus, nous souhaitons développer un outil d'aide à l'identification d'espèces d'arbres, c'est-à-dire qui va assigner un vecteur de caractéristiques à une liste d'espèces avec une mesure de confiance pour chaque espèce. Les arbres de décisions binaires n'offrent pas cette possibilité.

Les forêts d'arbres décisionnels (ou forêts aléatoires) se basent sur les arbres de décision. L'algorithme des forêts d'arbres décisionnels effectue un apprentissage sur de multiples arbres de décision entraînés sur des sous-ensembles de données légèrement différents. Il permet de corriger la sensibilité à l'ordre des prédicteurs (nœuds des arbres) d'un arbre de décision unique. Ce type d'algorithme a notamment été utilisé pour la classification de feuilles d'arbre [Cer+12] [Pac+14]. Malheureusement, l'algorithme des forêts aléatoires n'est pas supporté par des appareils avec une puissance de calcul et une taille mémoire limitée.

Il existe des classifieurs qui sont plus faciles à mettre en œuvre, comme la méthode des k plus proches voisins, utilisée pour la classification de feuilles [LP13], ou des mesures de distances [Cer+13a].

L'augmentation de la taille des bases de données a permis l'emploi de classifieurs dédiés à celle-ci, comme les SVM (Support Vector Machine) ou les réseaux de neurones. Lors des premiers challenges PlantCLEF, les SVM étaient souvent utilisés [Goë+13] [Goë+14]. En

effet, ceux-ci ont des temps de calcul faibles et donnent de bons résultats. Toutefois, depuis 2015, ce sont les réseaux de neurones qui sont les classifieurs les plus utilisés pour la reconnaissance des végétaux [GJB15]. Les réseaux de neurones, tels que le MLP (Multi-Layer Perceptron), donnent de très bons résultats, surtout sur des grandes bases de données avec un nombre important de classes (1000 classes pour le challenge PlantCLEF 2015), et avec des vecteurs caractéristiques de grande dimension. En revanche, leurs très bons résultats demandent un temps d'apprentissage considérable à cause du très grand nombre de données d'apprentissage nécessaires. Plus le réseau est grand (nombre de couches et nombre de neurones importants), plus temps de calcul pour la classification du vecteur de caractéristiques est long.

2.5 Les réseaux de neurones à convolutions (CNN)

Les MLP ne peuvent traiter des images de grandes dimensions à cause du nombre de connexions trop important entre les neurones. En effet, il y a autant de neurones sur la première couche (couche d'entrée) qu'il n'y a de pixels dans l'image, multiplié par le nombre de canaux, et chaque neurone est connecté à tous les neurones de la couche suivante. Un CNN (ou réseau neuronal convolutif, en anglais *Convolutional Neural Networks*) se compose d'une couche d'entrée et d'une couche de sortie, ainsi que de multiples couches cachées. Les couches cachées d'un CNN se composent généralement de couches convolutionnelles, de couches de pooling, de couches de correction. Les poids appris par les premières couches convolutionnelles permettent de construire automatiquement des noyaux de convolution pertinents alors que les couches intermédiaires sont chargées de regrouper ces réponses brutes en motifs visuels de haut niveau. Les dernières couches entièrement connectées fonctionnent plus traditionnellement comme n'importe quel MLP sur la représentation de l'image résultant des couches précédentes. Il a été récemment prouvé que les réseaux de neurones convolutionnels profonds (avec plus de trois couches) obtiennent les meilleurs résultats sur des ensembles de données de classification d'images à grande échelle comme ImageNet [KSH12].

Dans un premier temps, les réseaux de neurones étaient utilisés essentiellement dans le cadre de la classification sur la base de vecteurs descripteurs. Cependant, dans le domaine de l'identification automatique d'espèces végétales en traitement d'images, comme dans de nombreux domaines, une bascule s'est effectuée entre des descripteurs déterminés par l'expertise du domaine et l'utilisation des réseaux de neurones à convolution pour l'extraction de descripteurs. Plutôt que de créer des caractéristiques "à la main", ces méthodes basées sur des réseaux de neurones extraient directement de l'information sur le signal image. Le changement de méthode est très perceptible via les méthodes utilisées par les différentes équipes du challenge PlantCLEF depuis 2015 [GJB15 ; Jol+16b].

Les réseaux de neurones à convolution utilisés notamment pour l'identification de

végétaux sont AlexNet et GoogLeNet [Goë+14 ; GJB15]. Les caractéristiques pertinentes étant automatiquement détectées par les couches de convolution, il n'est pas nécessaire d'avoir une étape de segmentation avant de traiter l'image. Parfois même, les chercheurs regroupent certains types d'organes tels que les fleurs avec les fruits [GMC15].

Au vu du nombre important de paramètres à apprendre, les CNN requièrent un très grand nombre de données d'apprentissage. Il est à noter que, comme le nombre d'images de la base de données du challenge PlantCLEF de 2015 n'était pas assez conséquent pour obtenir de très bonnes performances (100000 images pour 1000 classes), les participants au challenge ont donc dupliqué les images en variant aléatoirement la luminosité et l'orientation [Cha+15], ou simplement en inversant horizontalement l'image [Cho15]. Le nombre d'images n'étant toujours pas suffisant pour des réseaux très profonds tels que GoogLeNet (22 couches), cela présente un risque de sur-apprentissage. Il est cependant possible de récupérer un modèle (une architecture composée de différentes couches) qui a auparavant été pré-entraîné sur une base de données plus grande telle que ImageNet [Rus+15]. Une fois que les paramètres du réseau ont convergé, une étape d'apprentissage supplémentaire est effectuée en utilisant les images de plantes pour affiner les poids du réseau (*fine tuning*) [Cha+15 ; Cho15 ; MYA16].

Le challenge PlantCLEF a été abandonné en 2017 car uniquement des solutions embarquant des CNN profonds étaient proposées. Il n'y avait donc plus de challenge, ni d'innovations [Jol+16a].

2.6 Cadre de travail

Pour rappel, le projet dans lequel s'inscrit cette thèse se base sur l'application pour smartphone d'identification d'espèces d'arbres via des photos de feuilles appelée *Folia*. L'objectif de cette thèse est d'enrichir cette application en améliorant l'identification des espèces grâce à des informations obtenues via les autres organes des arbres que sont les écorces, les fleurs ou les fruits, avec un regard botaniste, dans le but d'un objectif didactique d'enseignement de la botanique. Nous ne traiterons pas de l'identification d'espèces d'arbres via les feuilles car ce sujet a déjà fait l'objet d'une thèse [Cer13b] dont les méthodes sont implémentées dans l'application *Folia*. Différentes contraintes apparaissent donc :

— La contrainte principale est l'utilisation du smartphone. Bien que les puissances de calcul et de stockage augmentent chaque année, nous ne pouvons proposer une application trop lourde en mémoire ou nécessitant beaucoup de calculs. Il convient aussi de donner une réponse à l'utilisateur dans un temps raisonnable. Les activités d'identification végétale se déroulant généralement dans la nature, il se peut que l'utilisateur n'ait pas accès au réseau Internet. Nous devons en conséquence proscrire l'envoi de photos sur un serveur pour être traitées, et donc prévoir des al-

gorithmes d'extraction, de caractérisation et de classification des organes végétaux pouvant fonctionner sur smartphone.

— Le smartphone étant dans les mains d'un utilisateur non formé, la prise de photo des différents organes ne pourra pas être bien contrôlée. Il se peut que les photos ne soient pas de qualité optimale. Principalement, nous ne pouvons demander à l'utilisateur de prendre en photo les différents organes sur un fond uni. Les écorces, fruits et fleurs seront donc très majoritairement photographiés sur un arrière-plan naturel complexe, avec des variations de luminosité.

— L'utilisation d'une application sur téléphone portable permet cependant d'établir éventuellement une interactivité avec l'utilisateur. Celle-ci devra cependant être la plus limitée possible pour éviter de lui demander un trop grand nombre d'actions ou de décisions. Par exemple, il sera difficilement possible de demander à l'utilisateur novice en botanique de renseigner si l'organe qu'il a photographié est un fruit ou une fleur, d'après la conclusion du questionnaire présenté dans le Chapitre 1.

— Notons que l'application *Folia* a une vocation pédagogique. Il sera donc intéressant de fournir à l'utilisateur les caractéristiques botaniques de l'organe qu'il a pris en photo. Ces caractéristiques devront bien évidemment provenir de l'organe d'intérêt, donc une phase de segmentation semble essentielle.

— Enfin, la confusion qu'il existe entre les espèces, du fait de la variabilité intra-espèces et de la similarité extra-espèces, devra être considérée afin d'améliorer l'identification et informer l'utilisateur du risque de confusion.

2.6.1 Segmentation d'organes végétaux en milieu naturel complexe

Pour l'étape de segmentation, la méthode la plus robuste serait de demander à l'utilisateur de détourer lui-même l'organe végétal d'intérêt. Mais cette étape est vraiment contraignante pour l'utilisateur. Grâce à la possibilité d'interactions entre l'application et l'utilisateur, dans le cas de la prise de photo d'écorce, nous pouvons demander à l'utilisateur de se rapprocher suffisamment du tronc de façon à ce qu'il n'y ait que de la texture d'écorce dans la photo. Il pourra, d'autre part, pointer l'objet d'intérêt si nécessaire. Nous souhaitons en outre mettre en place un algorithme de segmentation performant, capable de segmenter tous types de fruits ou de fleurs. Dans la littérature, la segmentation des fleurs est étudiée pour les fleurs "classiques", c'est-à-dire que nous pouvons trouver chez le fleuriste et présentant des pétales colorés bien distincts. Or sur les arbres d'ornement, c'est-à-dire les arbres cultivés pour d'autres aspects que la production maraichère, les fleurs sont présentes en inflorescence, ou ne ressemblent pas à des fleurs pour un novice en botanique (voir Chapitre 1), comme dans la famille des *Betulaceae* (aulnes, bouleaux, noisetiers, qui possèdent des fleurs appelées chatons). De même pour les fruits, les algo-

rithmes de segmentation de la littérature se basent sur leur aspect coloré et leur forme circulaire, ce qui n'est pas forcément le cas pour les arbres des forêts et des parcs (Figure 2.21).

FIGURE 2.21 – Les fleurs (rangée du haut) ne ressemblent pas toujours à des fleurs, et les fruits (rangée du bas) ne sont pas toujours d'aspect circulaire et coloré.

De plus, l'utilisateur ne pourra pas indiquer à l'application s'il a photographié une fleur ou un fruit. La méthode de segmentation des fruits et des inflorescences devra donc pouvoir segmenter ces organes sans *a priori* de forme ou de couleur, indépendamment du type d'organe, et être robuste aux différents fonds complexes dans lesquels se trouveront les organes. Toutefois, les smartphones ont récemment commencé à intégrer deux caméras sur leur face arrière, ce qui augmente les possibilités d'acquisition et de traitement d'images. Nous proposerons donc un algorithme de segmentation exploitant ces possibilités.

2.6.2 Description et classification/fusion des informations basées botanique des différents organes

Nous avons passé en revue les différents algorithmes permettant de décrire les organes végétaux. Cependant, dans le but de créer une application didactique, nous souhaitons indiquer les caractéristiques botaniques des organes à l'utilisateur. Les descripteurs locaux tels que les SIFT et les SURF, ou les descripteurs fonctionnant telles des boîtes noires, ne portant pas d'informations sémantiques seront donc à proscrire. La récupération de caractéristiques sur les différents organes sera, de manière optimale, à coupler avec les informations requises par les clefs de détermination botaniques, malgré une prise de vue limitant cette extraction de traits botaniques.

Le projet ReVeRIES aspire à pouvoir identifier les espèces d'arbres présentes en France, ce qui représente environ 300 espèces. Il sera donc nécessaire de choisir des systèmes de classification pouvant classifier les différents paramètres de manière optimale malgré la grande variabilité qui existe au sein des classes et les grandes similitudes entre celles-ci, tout en étant embarqués sur smartphone.

2.6.3 Bases de données

Au cours des dernières années, le nombre de bases de données regroupant des photos de plantes a augmenté, de même que la taille de ces bases de données. Cela tient au fait que de plus en plus d'équipes de chercheurs dans le monde s'intéressent à la reconnaissance de végétaux. Souvent la récolte d'images fait appel à la science citoyenne et de nouveaux outils permettent de traiter de grandes bases de données.

Il existe des bases de données d'images spécialisées pour chaque organe. La plupart des bases de données contenant des images de feuilles ont été construites spécifiquement pour la vision par ordinateur. C'est le cas des bases de données Swedish Leaves, Smithsonian ou Flavia [Wu+07]. Les différentes classes contiennent un nombre d'images important et homogène entre les classes. Cependant, il y a peu de diversité au sein d'une même espèce : les feuilles prises en photo appartiennent généralement à quelques individus, sur une période de temps réduite, ce qui n'est pas représentatif de la diversité de l'espèce végétale. De plus, les feuilles sont prises à plat sur un fond blanc, ce qui permet d'avoir une segmentation parfaite et une très bonne détection des bords. Ces bases de données ne sont donc pas utilisables dans le contexte d'une application en extérieur, ou l'arrière-plan de l'image est complexe, et l'image prise par un utilisateur sur smartphone.

Il existe cependant des bases de données consacrées aux fleurs ornementales, la plus utilisée dans la littérature étant OxfordFlower [6]. Celle-ci est divisée en deux jeux de données : une avec 17 classes (80 images par classe) et une autre avec 102 classes (entre 40 et 258 images par classe). Les photos représentent des fleurs que l'on trouve au Royaume-Uni avec des échelles différentes et des variations de luminosité. Cependant, ce ne sont uniquement que des photos de plantes herbacées (donc pas d'arbres ou d'arbustes) et avec une forme de fleur comme se le représente le grand public, c'est-à-dire avec des pétales bien visibles. De la même façon, il existe des bases de données de fruits, telles que CIFAR-100 [7] qui possèdent une partie "fruits et légumes" ou FIDS30 [8]. Toutefois, les fruits présents dans ces bases de données sont les fruits que nous consommons, par exemple les pommes, bananes, fraises, melons, *etc.* Ces bases de données ne sont donc pas adaptées aux types de fruits que nous pouvons trouver sur les arbres "d'ornement" ou sur les essences forestières que nous ne cultivons pas pour leurs fruits. Pour les écorces, les bases de données sont construites par des chercheurs qui travaillent en général sur l'analyse de texture, pour leurs publications. Elles sont de taille très réduite, avec à peine une dizaine d'espèces différentes [WSG11].

La base de données correspondant le plus à notre problématique est issue de la partie PlantCLEF du challenge ImageCLEF [Goë+13]. Les photos de plantes sont issues du pro-

6. http ://www.robots.ox.ac.uk/ vgg/data/flowers/
7. https ://www.cs.toronto.edu/~kriz/cifar.html
8. http ://www.vicos.si/Downloads/FIDS30

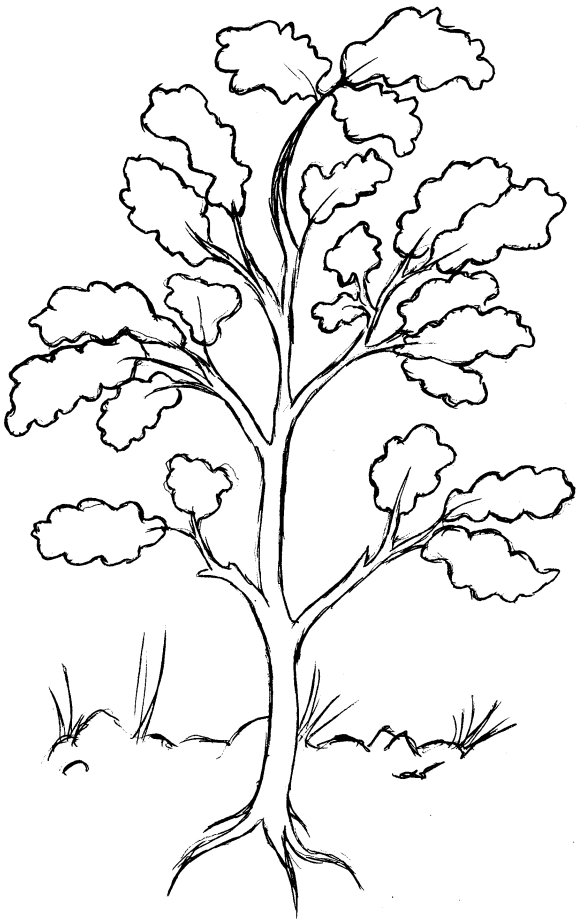
jet Pl@ntNet, un projet de recherche citoyenne s'appuyant sur le réseau Tela Botanica [9]. Les images sont prises et nommées par des citoyens, puis validées par des experts botanistes. La collecte d'images a commencé pendant l'été 2009 avec la récupération d'images de feuilles scannées. Puis une deuxième phase de collecte a eu lieu d'octobre 2010 jusqu'en mars 2011 pour obtenir de nouveaux scans de feuilles mais aussi des photos de feuilles sur un fond uni (dénommées *scan-like*). Ce n'est qu'à partir d'avril 2011 que la collecte d'images a été étendue aux autres organes des végétaux, c'est-à-dire les fleurs, les fruits, la tige (ou l'écorce), ainsi qu'aux photos de l'espèce dans son ensemble. Depuis 2014, la base de données LifeCLEF intègre aussi des images de branches.

Il existe une autre base de données utilisée par les applications de reconnaissance d'images se basant sur le big data et par les chercheurs traitant de l'identification de végétaux. Cette base est ImageNet [10]. Celle-ci n'est pas uniquement spécialisée dans les images d'organes de végétaux, mais possède néanmoins un nombre important d'images de plantes. Cependant, cette base de données contient beaucoup d'erreurs et des classes mal définies ou trop génériques (par exemple classe "chêne" qui regroupe des espèces différentes appartenant au genre *Quercus*). Les erreurs viennent du fait que les images ont été identifiées automatiquement, puis validées par au minimum deux utilisateurs, sans prendre en compte le degré de connaissance en botanique des utilisateurs comme c'est le cas pour la base de données PlantCLEF.

Pour toutes ces raisons, nous aurons dans le cadre de cette thèse à construire une base de données d'apprentissage et de tests adaptée à nos besoins.

9. http ://www.tela-botanica.org/
10. http ://image-net.org/

Chapitre 3
Reconnaissance d'écorces



Ce chapitre détaille les algorithmes mis en place afin de caractériser les structures des écorces dans le but d'identifier l'espèce d'un arbre à partir d'une image de son tronc. Dans une première partie nous exposons les différentes contraintes de la reconnaissance des écorces, d'un point de vue traitement d'images mais aussi botanique. Nous présentons alors notre approche pour extraire des caractéristiques à partir des photos. L'influence des différents vecteurs caractéristiques ainsi que leurs paramètres sont ensuite discutés dans une deuxième partie. Enfin, nous comparons notre méthode à un autre algorithme de reconnaissance de texture récemment publié et sur une autre base de données.

3.1 Cadre de travail

Nous avons vu, dans le Chapitre 1 traitant de la description botanique des différentes parties des végétaux, que la tige était l'organe le moins utilisé pour l'identification des plantes. Mais elle est régulièrement décrite dans les cas des espèces d'arbres, où l'on parle alors de tronc. Dans ce cas, c'est l'aspect texturé ou coloré de l'écorce (formée pour rappel d'un ensemble de cellules mortes) qui est exposé lorsqu'elle est caractéristique d'une espèce. Malgré sa faible utilisation et le peu de clefs de détermination se basant uniquement sur l'écorce pour l'identification d'espèces d'arbres, nous avons décidé de la prendre en compte dès le début car, contrairement aux fruits et aux fleurs portant une plus grande information botanique, celle-ci est facilement accessible et toujours présente sur l'arbre. Il existe de rares exceptions où l'écorce est absente du tronc si elle est broutée par les animaux sur les jeunes rameaux, ou pour le cas du chêne-liège (*Quercus suber*) dont l'écorce est enlevée par l'homme pour extraire le liège. Par ailleurs, l'écorce peut être très caractéristique pour certaines espèces d'arbres (bouleau, platane).

3.1.1 Évolution biologique de l'écorce

Pour rappel, les arbres possèdent une écorce du fait de la croissance secondaire de la tige, c'est-à-dire une croissance en largeur. Avec la croissance du diamètre du tronc, l'écorce du tronc au début lisse est obligée de se craqueler sous la poussée des nouvelles cellules. Ainsi, les jeunes arbres présentent généralement une écorce lisse, et les vieux arbres une écorce plus structurée dont la densité, l'orientation ou la largeur des craquelures dépendent de l'espèce à laquelle ils appartiennent (Figure 3.1).

La structure de l'écorce n'est pas uniquement corrélée à l'âge de l'arbre. Elle est aussi liée aux conditions environnementales (situation géographique, altitude, nature du terrain, ensoleillement, conditions hydrologiques, influence d'autres espèces végétales ou animales) qui affectent la croissance plus ou moins rapide du tronc de l'arbre. Il existe donc une large variabilité intra-classe. L'ouvrage de Jean-Denis Godet [God12], rédigé pour aider l'identification des espèces d'arbres d'Europe via les écorces, propose six photos d'écorces

Figure 3.1 – Exemples d'écorces de *Quercus ilex* à différents âges, d'abord lisses puis crevassées.

pour chaque espèce décrite à différents âges (juvénile, mature, âgé) dans différents environnements.

D'un autre côté, les écorces ne sont généralement pas utilisées pour l'identification des espèces car il existe une grande similarité inter-classe. En effet, des arbres appartenant à des clades botaniques différents peuvent présenter des écorces semblables (Figure 3.2).

Figure 3.2 – Similarité inter-classes entre les arbres *Acer platanoide*, *Populus nigra*, *Quercus nigra*, et *Tilia platyphyllios*.

La faible variabilité inter classe ainsi que la forte variabilité intra classe rendent plus ardue la tâche d'identification d'une espèce d'arbre via l'écorce.

Afin de caractériser les écorces, nous avons décidé de nous baser sur les critères pris en compte par les botanistes, d'une part pour réduire autant que possible la taille du vecteur descriptif, et d'autre part pour pouvoir indiquer au fur et à mesure du processus de reconnaissance les éléments caractéristiques de l'écorce photographiée. Plus précisément, la méthode simule la reconnaissance de l'écorce décrite par Michael Wojtech dans son livre [Woj11]. Dans cette approche, les écorces sont d'abord classées en différentes familles, comme le montre la Figure 3.3.

La première ligne contient des photos d'écorces caractéristiques pour chaque famille. La seconde ligne met en évidence les éléments distinctifs des différentes structures. Après avoir identifié le type d'écorce, les botanistes considèrent la couleur et d'autres éléments structurels pour identifier l'espèce d'arbre.

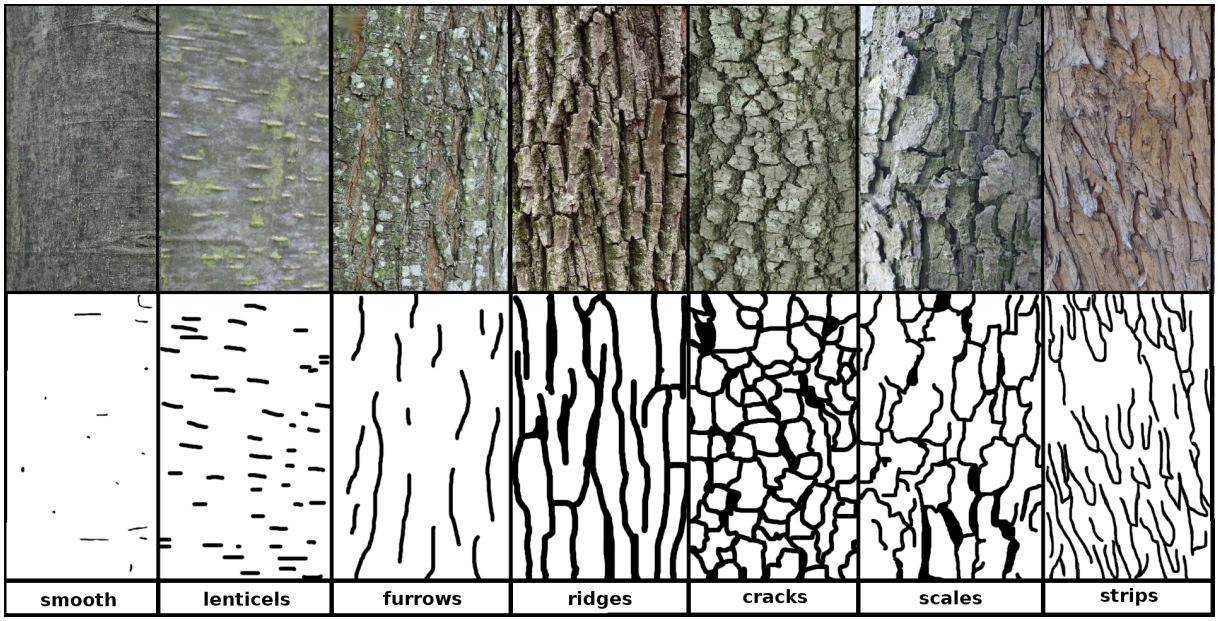


FIGURE 3.3 – Les différentes familles de structures d'écorce basées sur des critères visuels, avec des dessins mettant en évidence les structures distinctives des écorces.

3.1.2 Base de données utilisée

Pour tester l'efficacité des vecteurs de caractéristiques développés dans cette thèse, nous utilisons un sous-ensemble des images d'écorces de la base de données PlantCLEF. Nous avons gardé 101 classes parmi les espèces d'arbres sélectionnées lors du projet ANR ReVeS ayant des images d'écorces disponibles dans la base. Notre base de 101 espèces d'arbres est composée de 2559 images. Ces images ont été recadrées de manière à n'avoir que de la texture d'écorce dans l'image (Figure 3.4). En effet, nous avons supposé que dans notre cas, l'application pour smartphone peut guider l'utilisateur lors de la prise de vue, et ainsi indiquer à celui-ci de prendre la photo en mode portrait et de se rapprocher suffisamment du tronc afin de n'avoir que de l'écorce sur la photo. Il n'y a donc pas d'étape de segmentation.

FIGURE 3.4 – Exemple de recadrage d'une image d'écorce (image initiale à gauche et image recadrée à droite).

Pour rappel, les écorces ont été photographiées par des amateurs, sans consignes de prise de vue. Ainsi, il y a des variations de luminosité mais également des variations d'échelles, puisque les troncs n'ont pas tous été photographiés à la même distance. Par ailleurs, un inconvénient de cette base de données est que certaines espèces ont un grand nombre d'images d'écorces (plus de 100) et d'autres très peu (2 ou 3 éléments). En moyenne la population des classes est de 20 images (voir Figure 3.5). Ceci vient du fait que les citoyens alimentant la base PlantCLEF reconnaissent quelques espèces communes telles que le platane (*Platanus*) ou le chêne (*Quercus*) et vont fournir des images de ces espèces uniquement. Avoir des populations de classes déséquilibrées génère des difficultés pour entraîner un classifieur efficace.

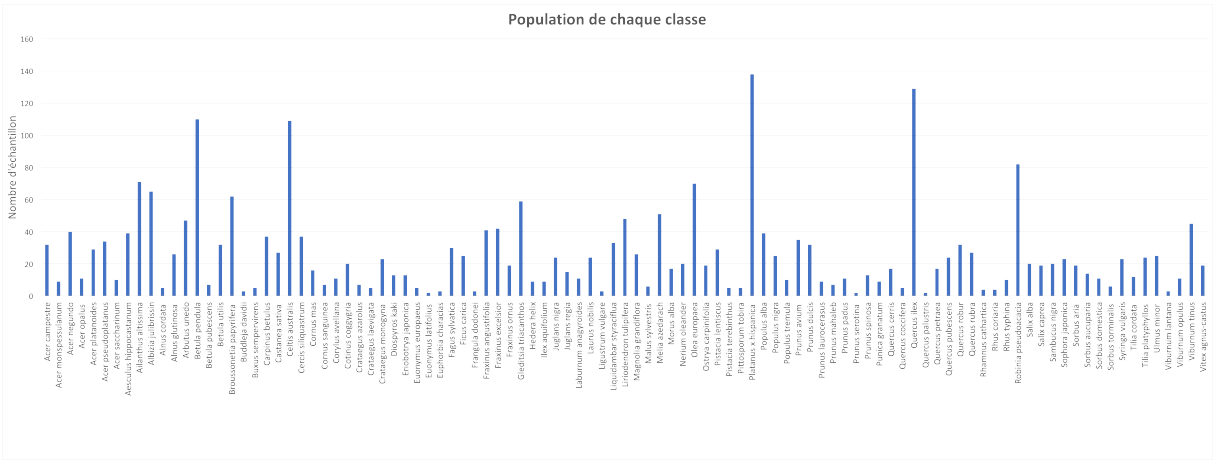


Figure 3.5 – Visualisation du nombre d'images de la base pour chaque espèce.

3.2 Extraction des caractéristiques

Cette partie décrit les différents vecteurs de caractéristiques créés de manière à intégrer de la sémantique, c'est-à-dire avoir une signification botanique. Ainsi, de même que les botanistes, nous souhaitons caractériser l'orientation de l'écorce, sa structure, et sa couleur.

3.2.1 Orientation principale de l'écorce

Nous pouvons voir sur la Figure 3.3 que les différentes familles d'écorces présentent des orientations différentes. Elles peuvent être lisses, ou présenter une structure horizontale (écorces présentant des lenticelles), ou verticale (écorces présentant des fissures, sillons, lanières) ou bien encore horizontale et verticale (écorces présentant des craquelures, écailles). Un outil efficace pour caractériser l'orientation et la fréquence d'une texture est le filtre de Gabor.

La fonction de Gabor est une fonction gaussienne modulée par une fonction sinusoïdale. La fonction de Gabor s'écrit de la manière suivante :

$$G(x, y, \theta, f) = exp\left(-\frac{1}{2}(\frac{x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2})\right) cos(2\pi f x_\theta) \tag{3.1}$$

$$avec \quad x_\theta = x \cdot cos(\theta) + y \cdot sin(\theta) \tag{3.2}$$

$$et \quad y_\theta = y \cdot cos(\theta) - x \cdot sin(\theta) \tag{3.3}$$

où θ est l'orientation de la sinusoïde, f sa fréquence, σ_x et σ_y les écarts-types de la gaussienne selon les axes des abscisses et des ordonnées. L'application de cette fonction de Gabor à un masque de convolution forme un filtre de Gabor (Figure 3.6).

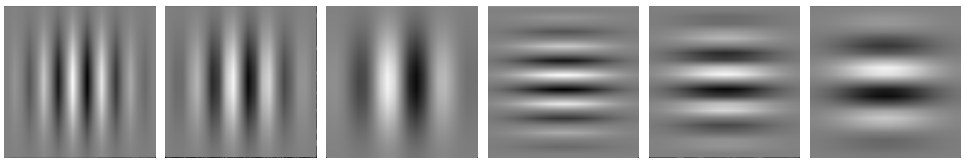


FIGURE 3.6 – Exemple de filtres de Gabor avec deux orientations θ et différentes valeurs de fréquences f.

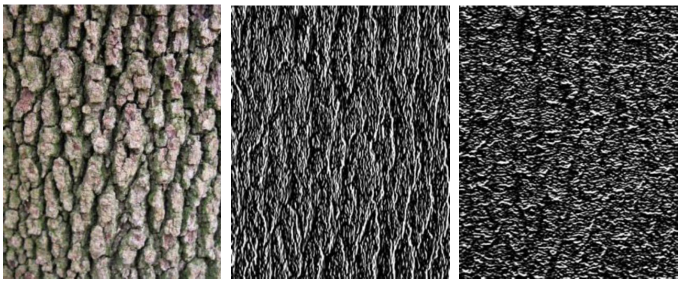
Le filtrage de Gabor consiste à convoluer une image I avec un masque de convolution M formé à partir du filtre de Gabor g.

$$g(I) = M * I \tag{3.4}$$

Le filtrage de Gabor permet d'isoler les contours d'une image ayant la même orientation θ que le filtre avec une épaisseur correspondant à la période $1/f$ de la fréquence de sinusoïde. Ce filtrage s'effectue sur des images en niveaux de gris. Nos images étant en couleur (R, G, B) la conversion en niveaux de gris G' est effectuée selon l'équation suivante :

$$G' = 0.299.R + 0.587.G + 0.114.B \tag{3.5}$$

Les filtres de Gabor ont notamment été utilisés en traitement d'images pour la classification des écorces d'arbres [CHC03 ; Hua+06a]. Selon le travail de Huang *et al.* [Hua+06a], seulement 6 orientations et 4 échelles de la sinusoïde sont suffisantes pour l'analyse des écorces. Cependant, dans le cas de notre application, la distance entre le smartphone et le tronc d'arbre peut varier. Il faut donc que la méthode soit indépendante de l'échelle. De plus, nous souhaitons déterminer dans une première étape si la structure de l'écorce est horizontale ou verticale, donc uniquement deux orientations ont été retenues (0° et 90°) (Figure 3.7).



(a) (b) (c)

Figure 3.7 – Résultats des filtrages de l'image initiale a) avec une orientation de 0° b) et 90° c).

Afin d'avoir un vecteur caractéristique de l'écorce indépendant de la distance de prise de vue, nous faisons la moyenne de quatre images filtrées par ondelettes de Gabor avec quatre fréquences sinusoïdales différentes. Les fréquences sont ajustées afin de mettre en évidence les structures (fissures et crevasses) horizontales et verticales de l'écorce. Le vecteur de caractéristiques $Gab(I)$, reflétant l'orientation de la structure de l'image I à l'aide de filtres de Gabor, se compose de la moyenne et de l'écart-type de l'image filtrée avec une orientation de 0°, σ_0 et μ_0, et de l'écart-type de l'image filtrée à 90°, σ_{90}. Nous avons envisagé d'inclure également la moyenne des images à 90°, μ_{90}, mais cette caractéristique est apparue comme non discriminante, nous avons donc décidé de l'exclure, pour finir avec un simple vecteur de caractéristiques de dimension 3 : $Gab(I) = (\sigma_0, \mu_0, \sigma_{90})$.

3.2.2 Description de la texture

Nous souhaitons maintenant décrire de manière plus précise la structure de l'écorce, et non seulement indiquer si elle a une structure verticale ou horizontale ; le but étant de pouvoir un peu mieux identifier les différentes familles d'écorces, c'est-à-dire faire ressortir les écailles, les lanières ou les craquelures de l'écorce, et caractériser leurs formes. Pour cela, les contours sont extraits à l'aide de l'algorithme de Canny [Can87]. C'est un algorithme largement utilisé car il a un faible taux d'erreur (bonne détection et position des contours) et donne une seule réponse par bord (pas de contour intérieur et de contour extérieur). Les étapes du détecteur de Canny sont les suivantes :

— Un filtrage gaussien est appliqué sur l'image afin de réduire le bruit. Dans notre cas, nous avons fixé de manière empirique la taille du noyau gaussien à 5 pixels avec un écart type de 0.3.

— L'image d'intensité du gradient est calculée grâce à l'opérateur de Sobel. L'opérateur utilise les noyaux de convolution F_H et F_V suivants pour calculer les dérivées

verticales et horizontales de chaque pixel.

$$F_H = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad F_V = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{3.6}$$

L'image du gradient horizontal est $G_H = F_H * I$ et l'image du gradient vertical est $G_V = F_V * I$. Puis la norme G et la direction θ du gradient général sont calculées de la manière suivante :

$$G = \sqrt{G_H^2 + G_V^2} \quad et \quad \theta = arctan(\frac{G_H}{G_V}) \tag{3.7}$$

— Une suppression des non-maxima est appliquée de manière à affiner les contours.

— Pour supprimer les pixels de bords qui sont causés par le bruit ou par des variations de luminosité, un seuillage à hystérésis est réalisé. Le seuillage à hystérésis possède un seuil bas et un seuil haut. Si le pixel de contour a une valeur de gradient inférieure au seuil bas, alors il est rejeté. Si sa valeur est supérieure au seuil haut, il est conservé comme contour. Si la valeur du gradient est comprise entre le seuil bas et le seuil haut, le pixel est accepté comme contour s'il est connecté à un pixel déjà accepté. Dans notre cas, le seuil bas a été fixé de manière empirique et le seuil haut est 2.5 fois ce seuil bas.

Les contours verticaux, formant l'image M_V, et les contours horizontaux, formant l'image M_H, sont calculés de manière indépendante. Les résultats sont montrés sur la Figure 3.8.

À partir de ces images M_V et M_H de contours, nous pouvons déjà distinguer les dissemblances entre les différentes familles d'écorce. Pour caractériser la répartition des contours, une grille est "placée" sur l'image de contours, et le nombre d'intersections entre les contours et la grille est compté. Autrement dit, à partir de la carte de contours verticaux M_V (respectivement horizontaux M_H), nous nous intéressons aux intersections entre ces contours et des lignes horizontales (respectivement verticales) de la grille. Ensuite, pour chaque ligne (respectivement colonne) de la grille, nous calculons le nombre d'intersections (figure 3.9). Cela génère un "mot vertical" W_V (respectivement "mot horizontal" W_H).

De manière plus formelle, soit l le nombre de lignes et c le nombre de colonnes dans l'image, avec λ_l la fréquence de la grille verticale et λ_c la fréquence de la grille horizontale. Plus les fréquences λ_l et λ_c sont élevées, plus la structure de l'écorce est décrite de manière précise. Le "mot vertical" W_V est une concaténation des intersections v_l avec l variant de 0 à l avec le pas de la grille λ_c.

$$v_l = \sum_{c=0}^{\lfloor \frac{c}{\lambda_c} \rfloor} \delta_{l,c} \tag{3.8}$$

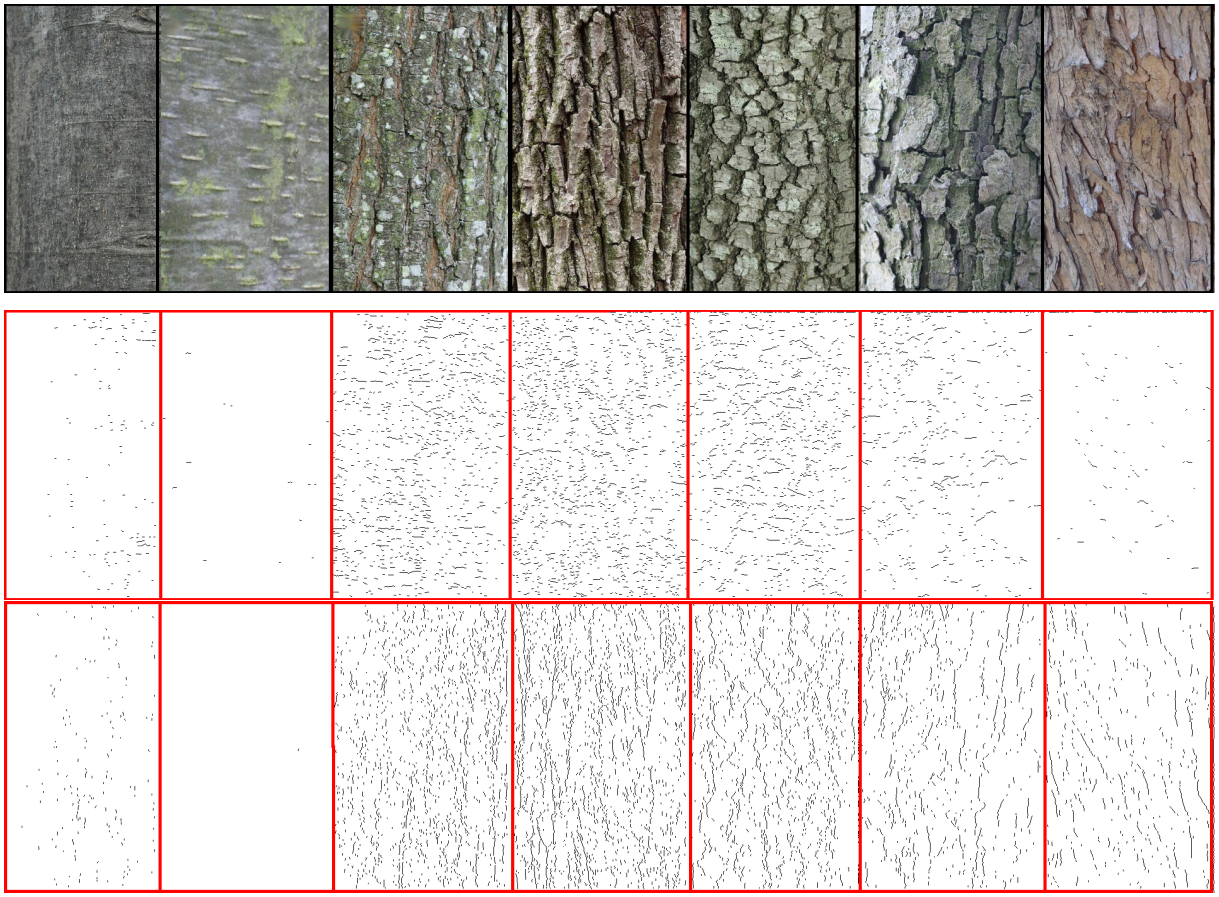


Figure 3.8 – Résultat de l'algorithme de Canny pour extraire les contours horizontaux et verticaux des images d'écorces types de chaque famille.

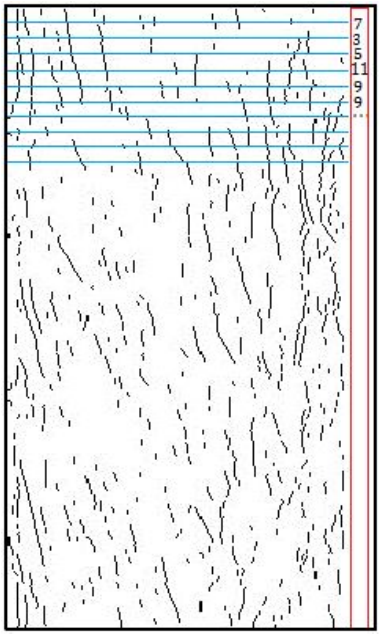


Figure 3.9 – Formation du "mot vertical" W_V.

$$\text{avec } \begin{cases} \delta_{l,c} = 1 & \text{si } M_V(\lambda_l, \lambda_c) = 1 \\ \delta_{l,c} = 0 & \text{sinon} \end{cases} \quad \text{et } V = \{v_l\}_{0 \leq l \leq \left\lfloor \frac{l}{\lambda_l} \right\rfloor} \tag{3.9}$$

M_H and M_V sont finalement caractérisés par deux mots, respectivement W_H and W_V, qui définissent la fréquence et la distribution des contours. Si un mot contient majoritairement des zéros, cela signifie que la structure de l'écorce dans l'orientation correspondante est assez lisse. Quand les fissures de l'écorce sont pratiquement ininterrompues, le mot vertical va être homogène. La Figure 3.10 donne des exemples de mots verticaux et de mots horizontaux pour deux exemples d'écorces. L'influence de la fréquence de la grille est discutée dans la partie 3.4.2. Pour finir, le mot vertical est normalisé par le nombre de colonnes de l'image d'écorce et le mot horizontal est normalisé par le nombre de lignes.

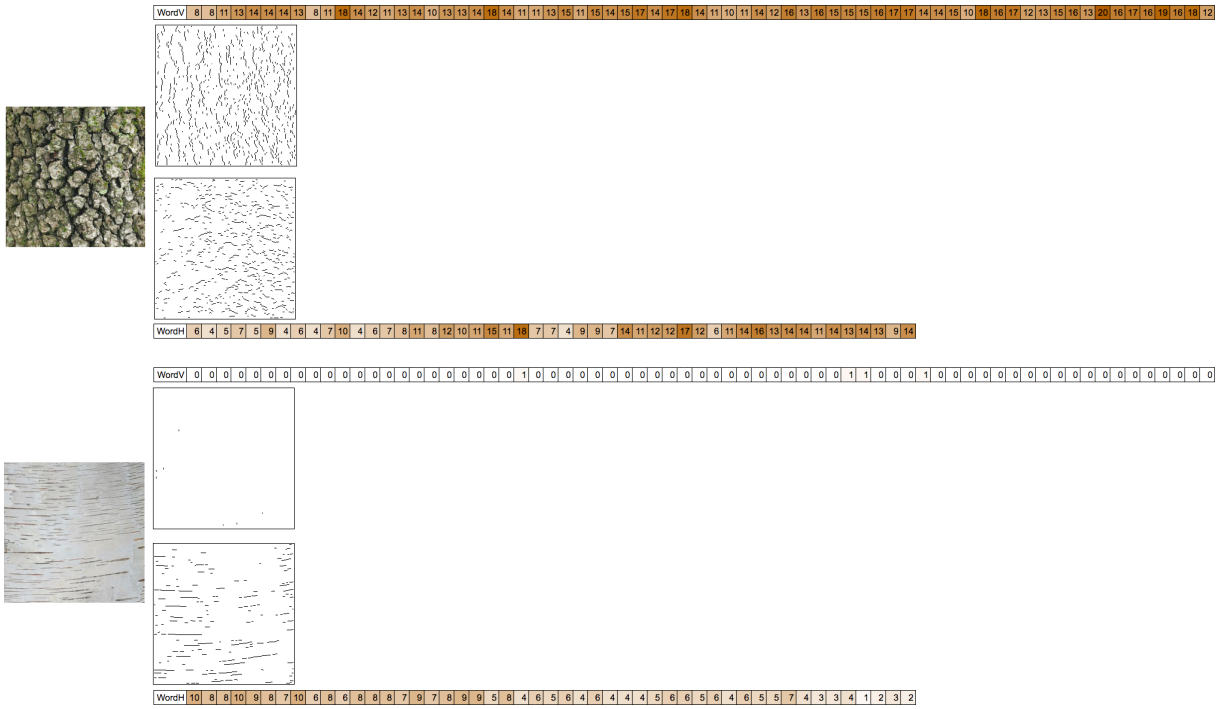


FIGURE 3.10 – Exemple de mots verticaux et horizontaux pour deux types d'écorce avant normalisation.

3.2.3 Caractérisation de la couleur de l'écorce

Un troisième levier permettant la reconnaissance des écorces d'arbres utilisé par les botanistes est la couleur. Lorsque la couleur des écorces est considérée, ce sont les espaces colorimétriques RGB et HSV qui sont généralement utilisés [Wan+04 ; Bak+13]. Wan *et al.* [Wan+04] appliquent des descripteurs statistiques de texture indépendamment sur chacun des canaux R, G et B des images d'écorces. Bakic *et al* [Bak+13] calculent des histogrammes RGB et HSV pondérés. Dans l'espace colorimétrique RGB, la couleur est codée sur 3 canaux de 256 valeurs, ce qui donne une taille importante de vecteur, d'autant plus que les histogrammes de chaque canal sont dépendants de la luminosité.

À noter que dans l'espace colorimétrique L*a*b*, la couleur est caractérisée par les deux canaux a* et b* (le canal L correspondant à la luminosité). Le canal a* code l'oppo-

sition rouge/vert et le canal b* l'opposition jaune/bleu. La plupart des arbres ayant des écorces marrons, vertes ou grises, le canal a* semblait intéressant à considérer, mais la teinte jaune reste influente dans la caractérisation de la couleur des écorces. Nous avons donc choisi de caractériser la couleur de l'écorce avec l'histogramme du canal H (la teinte) de l'espace colorimétrique HSV. Dans cet espace, la couleur (la teinte) est codée sur un seul canal et de manière indépendante à la luminosité (Équation 3.10). En effet, nous avons :

$$V = max(R, G, B)$$

$$S = \begin{cases} \frac{V - min(R,G,B)}{V} & si \quad V \neq 0 \\ 0 & sinon \end{cases} \tag{3.10}$$

$$H = \begin{cases} 60(G-B)/(V - min(R,G,B)) & si \quad V = R \\ 120 + 60(B-R)/(V - min(R,G,B)) & si \quad V = G \\ 240 + 60(R-G)/(V - min(R,G,B)) & si \quad V = B \end{cases}$$

La saturation (S) est l'intensité de la couleur. Plus la saturation d'une couleur est faible, plus l'image sera « grisée » et plus elle apparaîtra fade. La valeur (V) est la "brillance" de la couleur. Plus la valeur est faible, plus la couleur est sombre. La teinte (H) quant à elle est codée selon l'angle qui lui correspond dans le cercle chromatique (Figure 3.11).

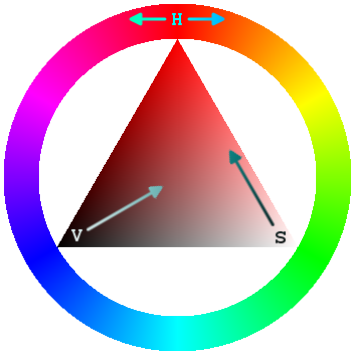
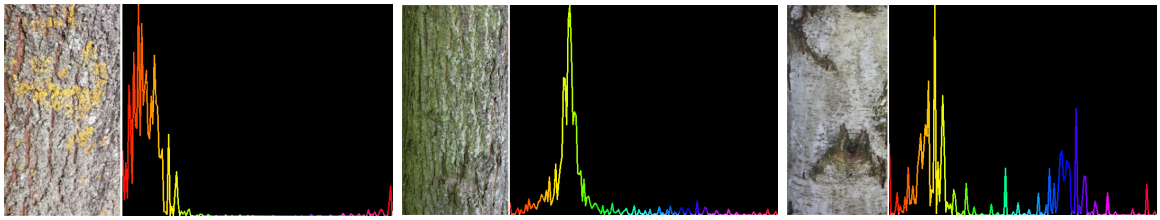


Figure 3.11 – Représentation de l'espace colorimétrique HSV.

La couleur de l'écorce est donc caractérisée par l'histogramme du canal H. La taille de cet histogramme H doit être de 360 bins. Cependant, cette valeur de 360 ne peut pas être stockée sur un seul octet. Pour des raisons de calcul logiciel, il est plus facile de transposer cette plage [0 ;359] dans la plage [0 ;179] que dans la plage [0 ;255]. Ainsi, la taille du vecteur est de 180 bins. Nous supposons que la perte d'un niveau de teinte sur deux n'est pas discriminante pour la description des couleurs des écorces. L'influence de la taille de l'histogramme de couleurs est cependant discutée dans la partie 3.4.3. La Figure 3.12 montre des exemples d'histogramme H pour certaines images d'écorces.



(a) *Acer campestre* 1 (b) *Acer campestre* 2 (c) *Betula pendula*

FIGURE 3.12 – Exemples d'images d'écorces et leur histogramme associé.

Pour résumer, notre vecteur caractéristique est composé du vecteur de Gabor, du vecteur de couleur, et des mots verticaux et horizontaux.

3.3 Classification

Cette partie présente le système de classification mis en place pour la reconnaissance d'écorces, ainsi que les résultats de classification des vecteurs d'écorce détaillés dans les parties précédentes.

Pour se rapprocher des clefs botaniques, une possibilité de reconnaissance est l'utilisation d'un arbre de décision. Le problème de l'utilisation des arbres de décision est que si l'algorithme se trompe lors d'un embranchement, le résultat de classification ne sera pas bon, sans pouvoir prendre en compte les autres caractéristiques. Par ailleurs, les écorces d'arbres peuvent présenter des caractéristiques différentes au cours de leur vie. Or, l'absence de détection d'un élément caractéristique (par exemple des lenticelles) ne doit pas gêner la classification. La littérature récente montre une grande utilisation des réseaux de neurones pour la classification [Cha+15]. Cependant, ces classifieurs nécessitent un grand nombre de données d'apprentissage pour obtenir des résultats de classification corrects, ce qui n'est malheureusement pas notre cas. Le choix s'est donc orienté sur un SVM (Support Vector Machine) ayant de bonnes performances sur des bases de données de faibles échantillons avec des temps de calcul faibles, ce qui est intéressant pour une utilisation sur smartphone. Avec un vecteur caractéristique de taille importante, une séparation linéaire est insuffisante. Une séparation polynômiale demanderait un degré de polynôme élevé difficile à supporter par un smartphone. Un noyau de séparation RBF (Radial Basis Function) [Vap99] a donc été choisi. Celui-ci permet de séparer des vecteurs dans de grandes dimensions.

3.3.1 SVM un contre tous

Afin d'avoir un score de probabilité d'appartenance à chaque espèce de manière à les ordonner, il a été mis en place la méthode de "SVM un contre tous". Ainsi, pour chaque espèce, nous créons un SVM avec deux labels : le label de l'espèce S, et le label de toutes les

autres espèces $non - S$. Le vecteur caractéristique de l'écorce inconnue est testé avec tous les SVM binaires de toutes les espèces. Nous gardons ensuite les N espèces correspondant aux N plus grands scores de probabilité d'appartenance. Nous obtenons donc une liste de N espèces probables pour une image d'écorce inconnue.

L'apprentissage est effectué sur la moitié de la base de données de 101 classes d'écorces, et les tests sont effectués sur la seconde moitié. Le vecteur caractéristique est composé du vecteur de Gabor, du vecteur de couleur, et des mots vertical et horizontal. Dans la continuité de l'application *Folia*, nous souhaitons retourner une liste comportant les dix espèces les plus probables pour une écorce inconnue. Le graphique de la Figure 3.13 présente les taux de classification pour N allant de 1 à 10. Si nous prenons $N = 2$, le taux de classification correspond à la probabilité d'avoir la bonne espèce parmi les espèces ayant les deux probabilités les plus élevées.

Les résultats de classification sont présentés sur la Figure 3.13.

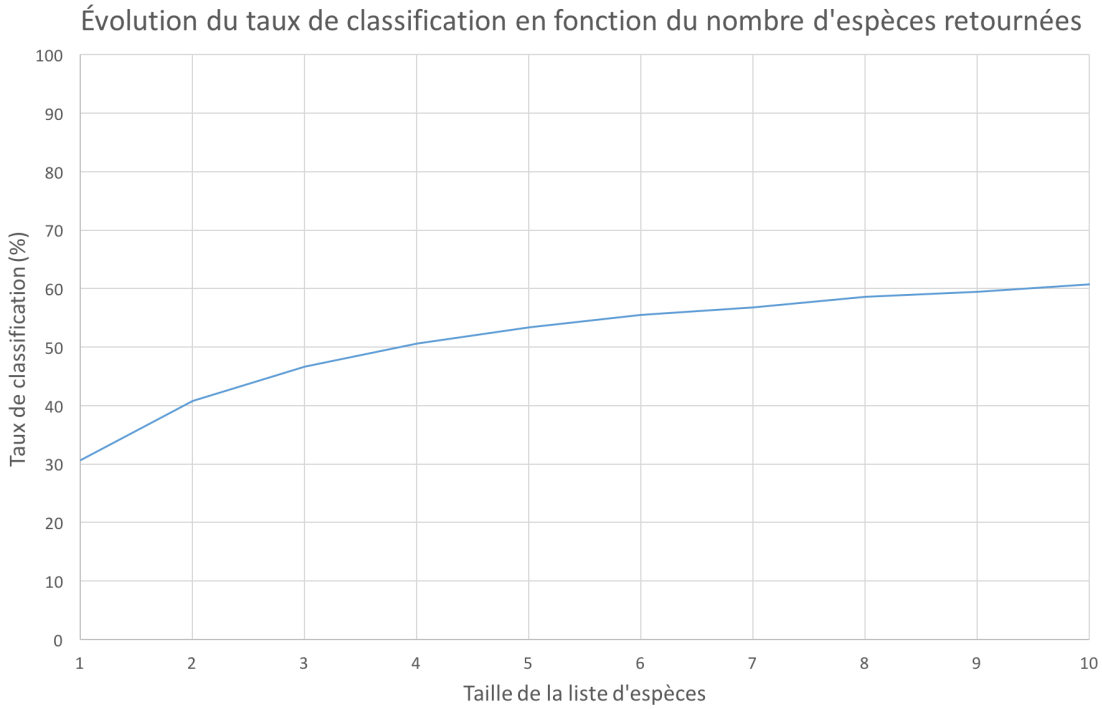


Figure 3.13 – Résultats de classification pour une base de 101 écorces d'arbres.

Avec ce système de classification, le taux de reconnaissance est 31,73% pour une espèce retournée. Bien évidemment, le taux de classification augmente avec le nombre d'espèces retournées. Il augmente de 10% lorsque deux espèces sont proposées. Nous avons 50% de chance d'avoir la bonne espèce dans une liste de 4 espèces possibles. La progression diminue ensuite, pour à la fin avoir 62,07% de chance d'avoir la bonne espèce dans une liste de 10 propositions.

3.3.2 Classifeur Perceptron Multi-couches (MLP)

Bien que le nombre d'échantillons soit probablement insuffisant, les résultats de classification avec un MLP sur la base de 101 écorces sont présentés dans cette partie.

Pour rappel le réseau multi-couche est un réseau de neurones composé au minimum de trois couches de neurones. La première couche (couche d'entrée) possède autant de neurones que la dimension du vecteur caractéristique. La dernière couche (couche de sortie) contient autant de neurones que de classes. Le nombre de couches cachées (couches entre les couches d'entrée et de sortie), ainsi que le nombre de neurones sur ces couches cachées dépendent du problème de classification. Il faut cependant noter que plus le nombre de neurones et de couches est élevé, plus il y a de poids et de paramètres à calculer, et plus de données d'apprentissage sont nécessaires. Un conseil est cependant d'avoir un nombre de neurones sur la couche cachée compris entre le nombre de neurones d'entrée et le nombre de neurones de sortie.

Nous avons testé un réseau de neurones composé de 3 couches (1 couche cachée), et un réseau de neurones composé de 4 couches (2 couches cachées). Le nombre de neurones sur les couches cachées varie entre 70 et 300 (Figure 3.14 pour le MLP avec une seule couche cachée et Figure 3.15 pour le MLP avec deux couches cachées).

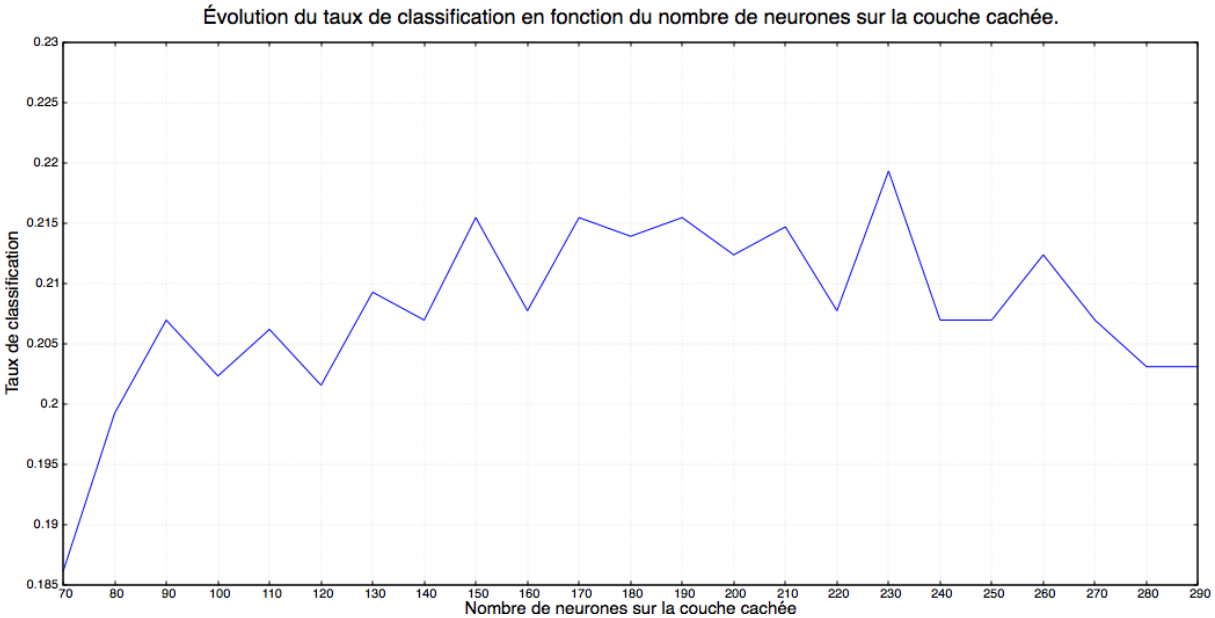


FIGURE 3.14 – Évolution du taux de classification en fonction du nombre de neurones sur la couche cachée pour un réseau de neurones composé de trois couches.

Dans un premier temps, nous pouvons remarquer que les taux de classification obtenus avec un réseau de neurones à 3 couches ou 4 couches sont inférieurs aux taux de classification obtenu avec le classifieur SVM (au maximum 21,93% et 26,64% respectivement, contre 31,73% avec un classifieur SVM). En effet, d'une part, une base de données de 101 classes constituées en moyenne de 20 images par classe n'est pas idéale pour régler tous

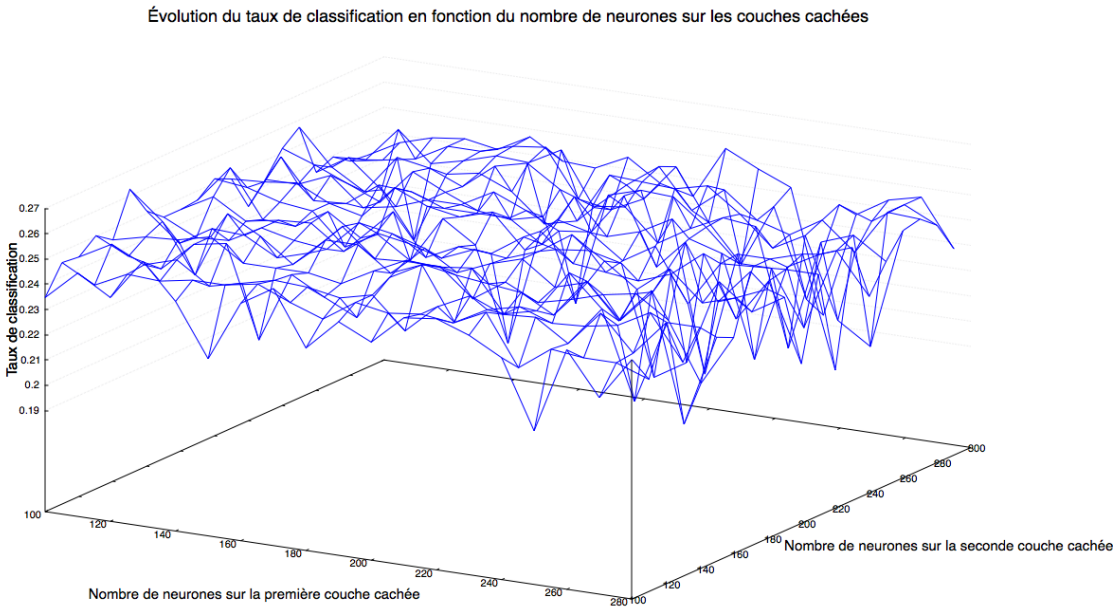


Figure 3.15 – Évolution du taux de classification en fonction du nombre de neurones sur les deux couches cachées pour un réseau de neurones composé de quatre couches.

les paramètres du réseau de neurones. D'autre part, nous pouvons voir que le nombre de neurones sur la ou les couches cachées influencent les taux de classification. L'évolution est assez faible pour le MLP avec une seule couche cachée mais plus importante pour le MLP avec deux couches cachées. Les MLP ne sont donc pas des classifeurs adaptés à notre problème de classification.

3.4 Influence des différents descripteurs

3.4.1 Influence du vecteur Gabor

Comme vu précédemment, le vecteur caractéristique des écorces est composé du vecteur de Gabor (Gab), du canal H de l'espace colorimétrique HSV (H), d'un mot vertical W_V et d'un mot horizontal W_H. Cependant, nous pouvons penser que les mots verticaux et horizontaux portent aussi l'information extraite par le filtrage de Gabor. Sur le même jeu de données que précédemment, nous regardons les taux de classification pour les 101 écorces d'arbres avec le vecteur caractéristique $V = (G, H, W_V, W_H)$, et le taux de classification avec le même vecteur caractéristique mais sans la composante issue de Gabor $V' = (H, W_V, W_H)$. Les résultats sont présentés sur la Figure 3.16.

La courbe bleue représente l'évolution du taux de classification des écorces avec le vecteur complet V, tandis que la courbe verte représente l'évolution du taux de classification sans l'information extraite par le filtrage de Gabor, soit V'. Nous pouvons remarquer que ces taux de classification sont bien moins bons quand nous ne prenons plus en compte

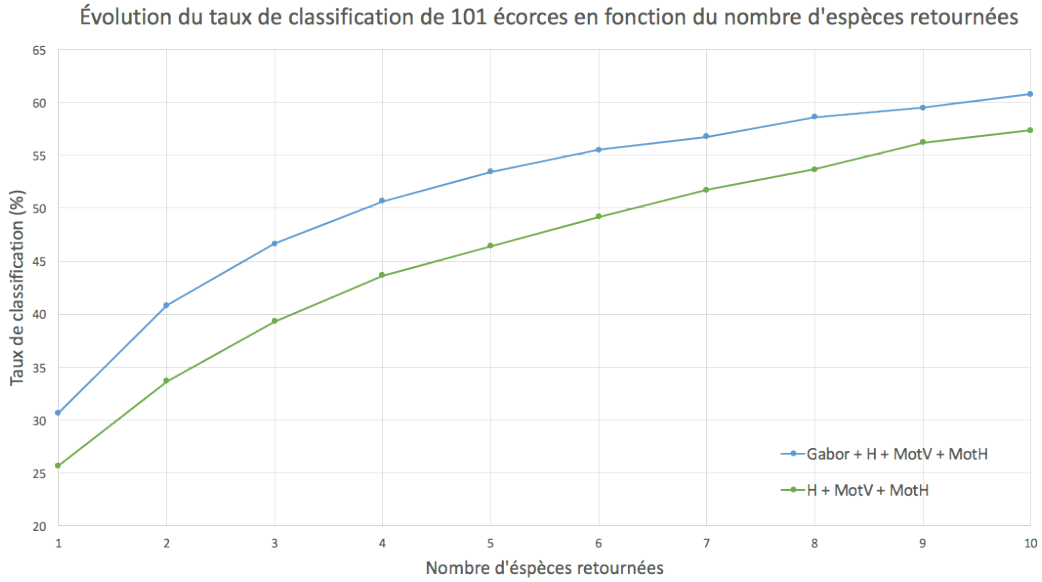


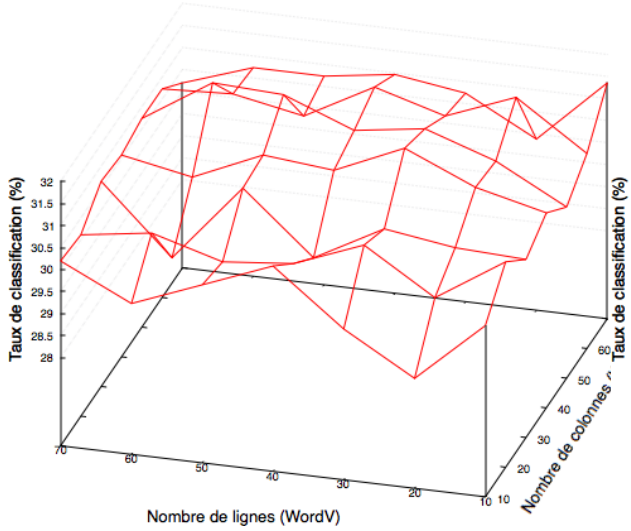
FIGURE 3.16 – Évolution du taux de classification des écorces selon la prise en compte ou non du vecteur caractéristique de Gabor.

le vecteur Gabor. En effet, nous perdons en moyenne 5% de bonne classification. De ces résultats, nous pouvons voir que les informations extraites par le filtrage de Gabor ne sont pas incluses dans le mot vertical ou le mot horizontal, et que sa prise en compte augmente la bonne classification des écorces.

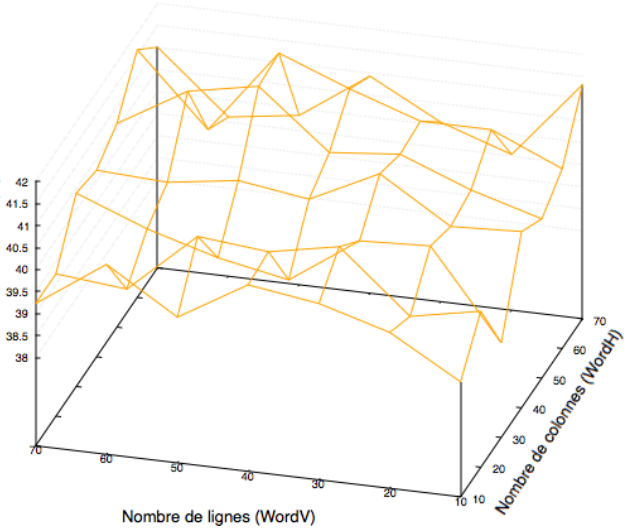
3.4.2 Influence des périodes du quadrillage pour l'extraction du "mot vertical" et du "mot horizontal"

Sur la base de données de 101 espèces, nous calculons les taux de classification en faisant varier le nombre de colonnes et de lignes du quadrillage permettant d'obtenir respectivement W_V et W_H. La Figure 3.17 montre l'évolution du taux de classification en fonction du nombre de lignes et de colonnes du quadrillage, pour une seule espèce retournée (Top 1), pour une liste de 2 espèces retournées (Top 2), *etc.*, jusqu'à une liste de 10 espèces retournées (Top 10).

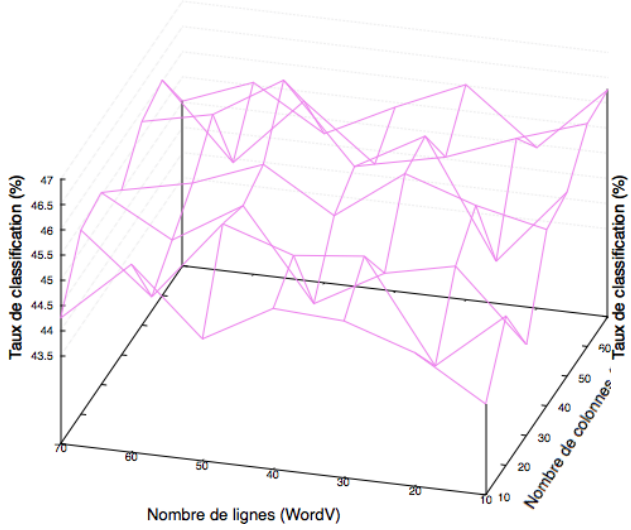
Le nombre de colonnes est limité à 70 car la définition de certaines images de la base de données n'est pas suffisante pour augmenter plus. Nous pouvons voir qu'en moyenne les meilleurs taux sont obtenus quand le nombre de lignes est égal à 50 et que le nombre de colonnes est de 70. Nous pouvons remarquer qu'au-dessus de 45 lignes, le taux de classification entame une décroissance. Plus le quadrillage est large (nombre de colonnes et de lignes faible), plus les taux de classification diminuent. Par ailleurs, nous pouvons noter que la résolution du quadrillage horizontal (nombre de lignes) pour la construction du mot vertical est plus élevée. Ceci vient du fait que les photos d'écorces sont en mode portrait et donc que la définition des images est plus grande en hauteur qu'en largeur.



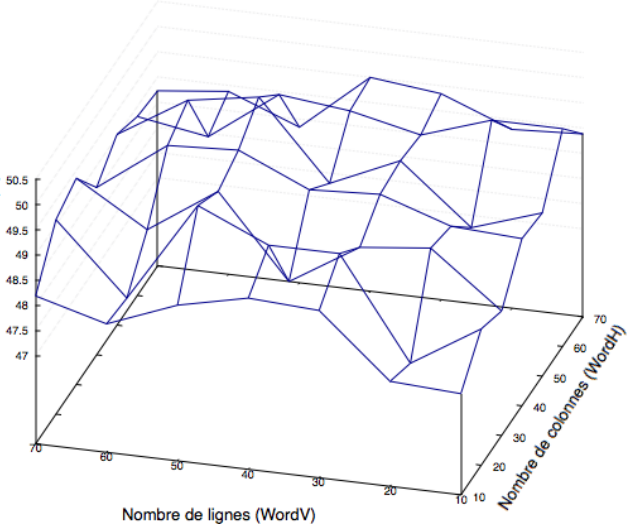
(a) Top 1



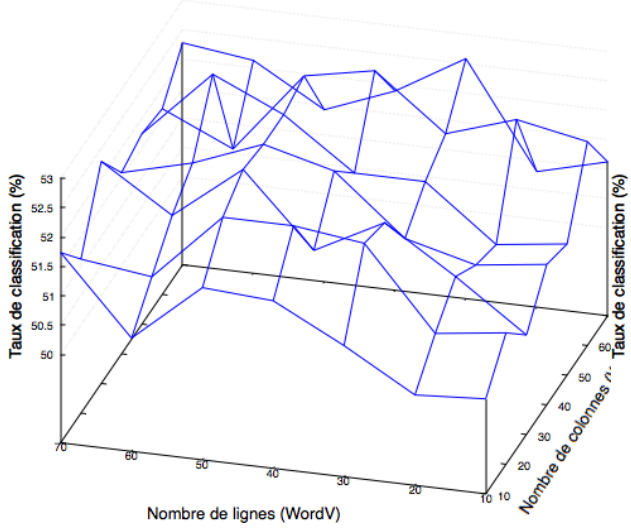
(b) Top 2



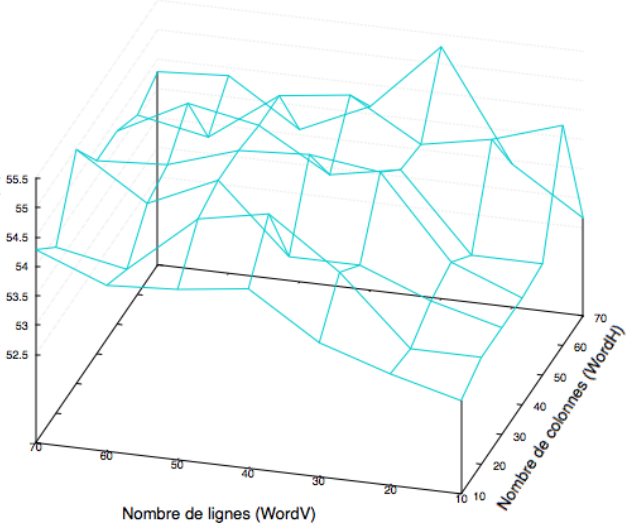
(c) Top 3



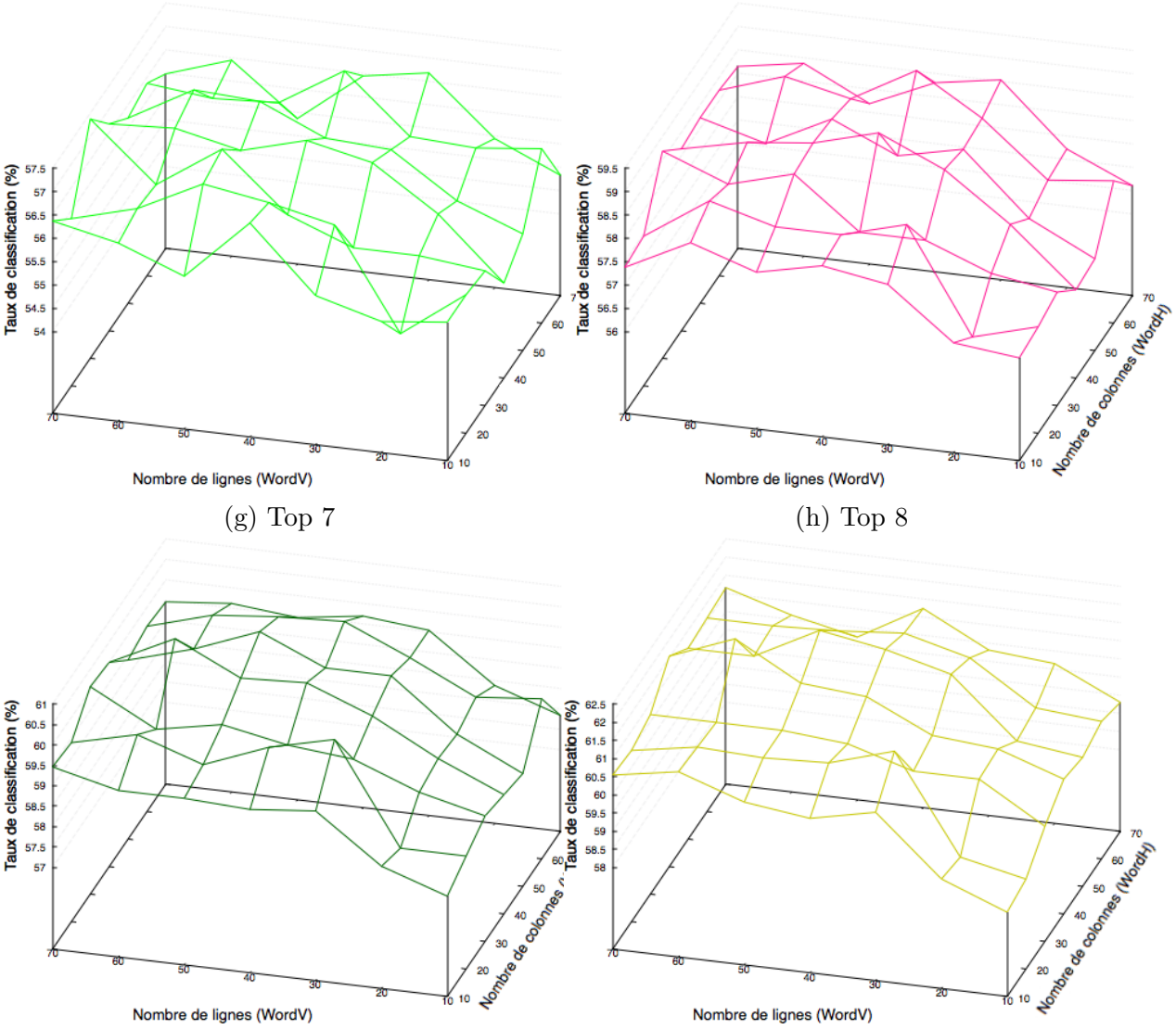
(d) Top 4



(e) Top 5



(f) Top 6



(g) Top 7 (h) Top 8

(i) Top 9 (j) Top 10

FIGURE 3.17 – Évolution du taux de classification en fonction du nombre de lignes et de colonnes du quadrillage pour la formation du mot vertical et du mot horizontal.

3.4.3 Réduction de l'histogramme

Notons que dans l'espace colorimétrique HSV toutes les couleurs ne sont peut-être pas utiles pour la classification des écorces d'arbres, particulièrement les teintes bleues ou violettes. Dans l'objectif de réduire au maximum le vecteur de caractéristiques des écorces, nous proposons de fusionner les bins de l'histogramme de couleurs les moins utilisés qui représentent les couleurs peu probables des écorces. Sur la base de données d'apprentissage, nous calculons tous les histogrammes H des images d'écorces. Pour déterminer les bins qui sont les plus utilisés, nous sommons ces histogrammes. Nous obtenons ainsi un histogramme "sommé" (Figure 3.18). Nous pouvons déjà voir que la majorité des écorces a des teintes rouges, oranges, jaunes, et vertes.

À partir de cet histogramme "sommé", nous fusionnons itérativement le bin de plus

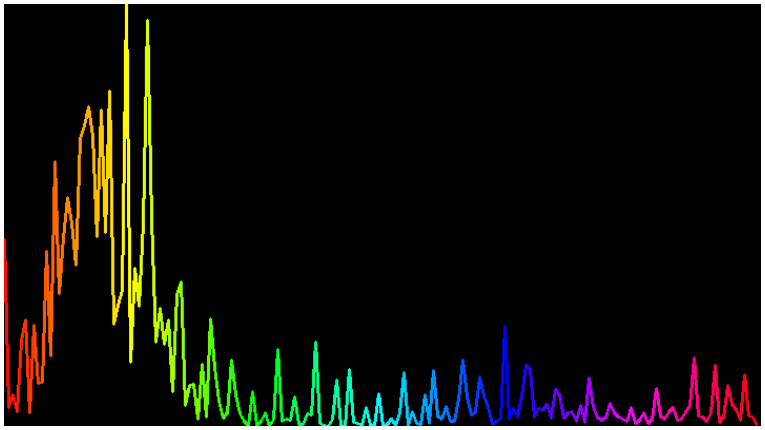


Figure 3.18 – Histogramme sommé des canaux H pour les écorces de la base de données d'apprentissage.

faible population avec son voisin le plus faible. Une attention a été portée au fait que le canal H de l'espace colorimétrique HSV est circulaire, et donc que les bins à chaque extrémité sont voisins. Une matrice de correspondance est aussi créée à chaque fusion de bin pour savoir quels sont les bins agrégés. Cette matrice de correspondance permet ensuite de calculer l'histogramme réduit en sommant les populations des bins agrégés. Pour savoir jusqu'à quel point nous pouvons quantifier l'histogramme, nous classifions la base de données de test d'écorce en réduisant de plus en plus la taille de l'histogramme de couleurs. La Figure 3.19 représente l'évolution du taux de classification global (Top 1) des images d'écorces classifiées uniquement avec le vecteur caractéristique de couleur en fonction de la taille de ce vecteur. Sachant que ce taux peut être influencé par le fait que certaines espèces possèdent plus de données de test que d'autres, l'évolution du taux de classification moyen par espèce est donnée sur la Figure 3.20.

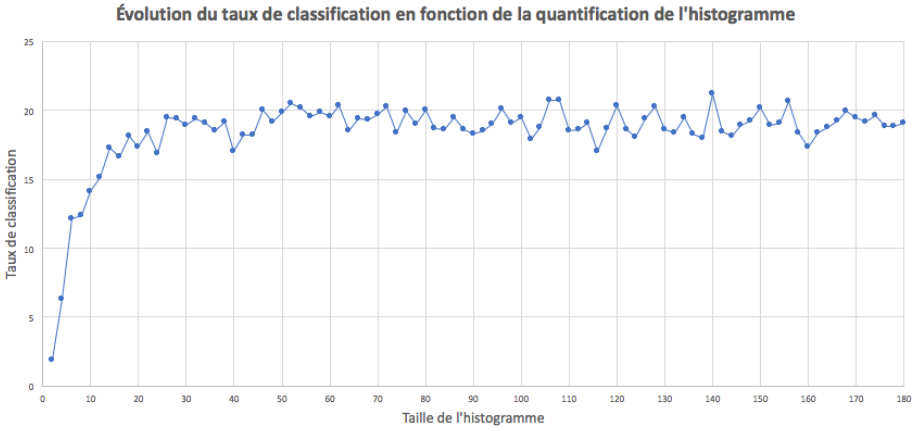


Figure 3.19 – Évolution du taux de classification global (Top 1) des images d'écorces classifiées uniquement avec le vecteur caractéristique de couleur en fonction de la taille de ce vecteur.

Malgré de nombreuses variations du taux de classification des espèces lors de la quantification de l'histogramme, nous pouvons voir que le taux reste constant avant de chuter

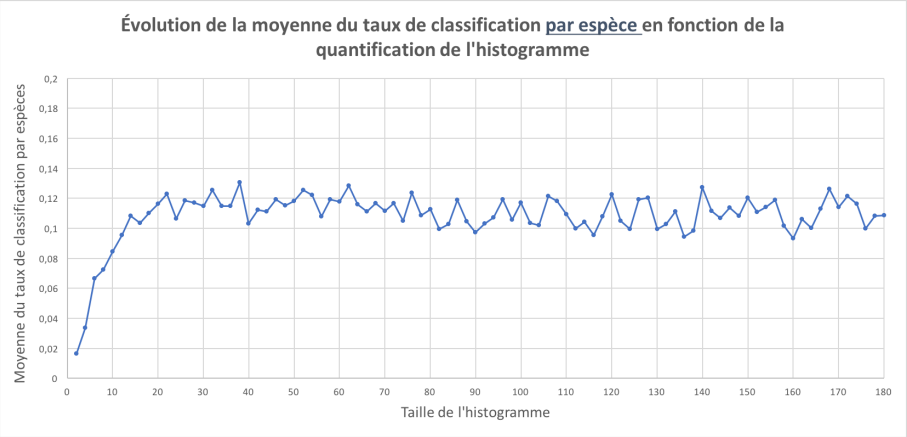


FIGURE 3.20 – Évolution du taux de classification moyen par espèce des images d'écorces classifiées uniquement avec le vecteur caractéristique de couleur en fonction de la taille de ce vecteur.

quand l'histogramme est réduit à 20 bins (Figure 3.19). En étudiant les taux de reconnaissance moyens pas espèces (Figure 3.20) nous pouvons remarquer que le taux augmente globalement lorsque la taille de l'histogramme est réduite. Nous pouvons en déduire qu'une définition plus précise de la couleur, avec un nombre de bin plus important, rajoute de la confusion lors de la classification.

Le vecteur descriptif de l'écorce est aussi composé des informations issues du filtrage de Gabor, du mot vertical et du mot horizontal. La Figure 3.21 présente l'évolution des taux de classification selon le nombre d'espèces prédites en fonction de la taille de l'histogramme H pour le descripteur complet.

Ici aussi, il est difficile de déterminer une taille d'histogramme optimale. Nous pouvons remarquer que plus la bonne espèce est prédite en dernier (elle est prédite en neuvième ou dixième position), plus la quantification de l'histogramme réduit sa prédiction. Par exemple, l'histogramme peut être réduit jusqu'à 20 bins pour le Top 1 en donnant les mêmes taux de classification sans réduction, alors qu'il ne peut être réduit que jusqu'à 78 pour le Top 10 avant de décroître fortement. Pour cette base de données, la question est donc de savoir s'il est préférable d'avoir le plus de chance que la bonne espèce se trouve dans les premières espèces proposées, ou de l'avoir au moins dans la liste de 10 espèces.

Sachant que le nombre d'échantillon de test n'est pas le même pour toutes les classes, nous proposons de fixer la taille de l'histogramme de couleurs en fonction du taux de classification moyen par espèce. La Figure 3.22 montre l'évolution de cette moyenne en fonction de la taille de l'histogramme H réduit.

De même que pour les précédents graphes, le taux de classification reste globalement constant avant de décroître fortement quand le vecteur n'est plus composé que d'une dizaine de bins. Nous pouvons voir que ce décrochage s'effectue lorsque la taille de l'histogramme est en dessous de 20 bins. En prenant en compte les informations des graphes précédents, nous avons fixé le nombre de bins à 30 pour caractériser la couleur de l'écorce.

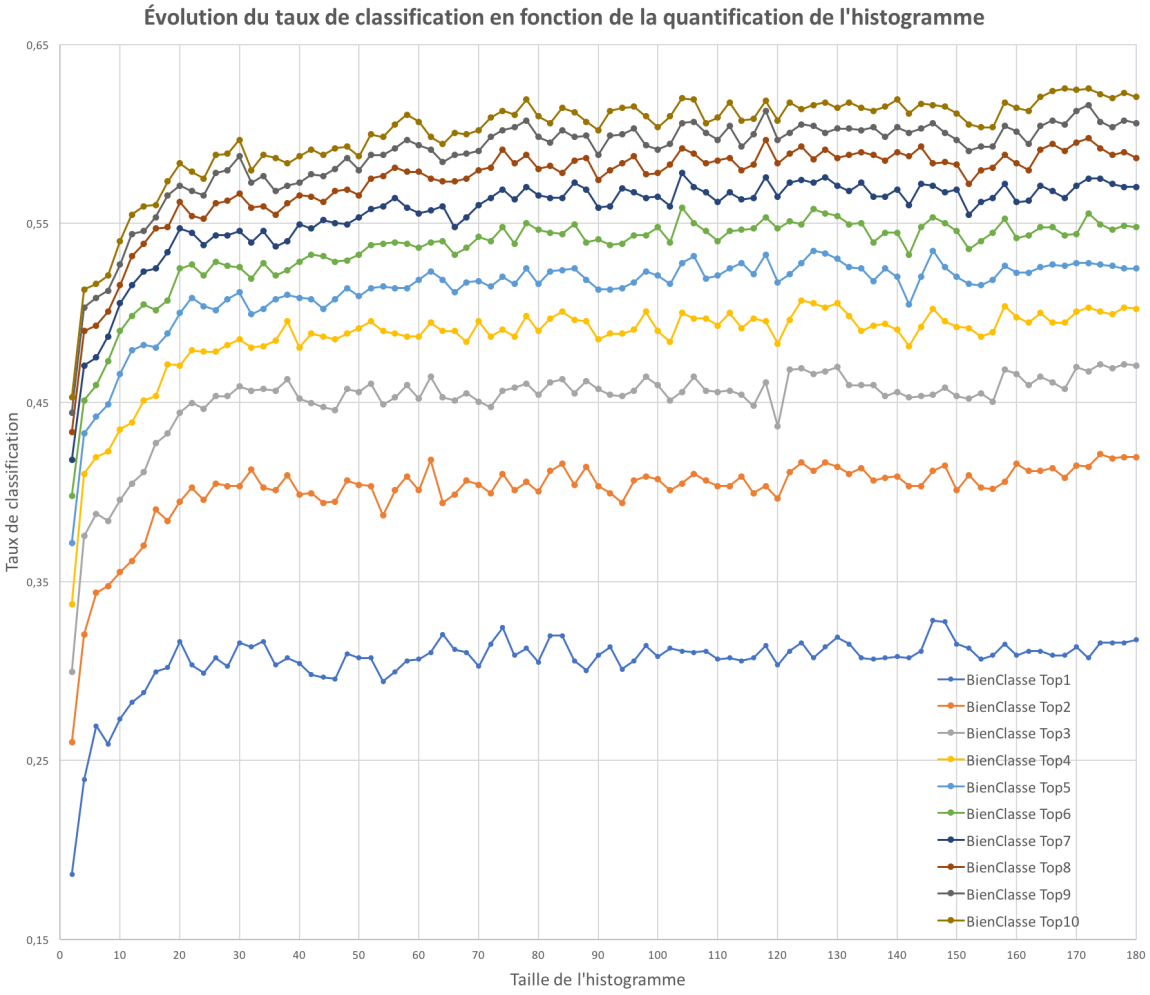


Figure 3.21 – Évolution des taux de classification globaux des images d'écorces classifiées avec le vecteur caractéristique complet en fonction de la taille de ce vecteur.

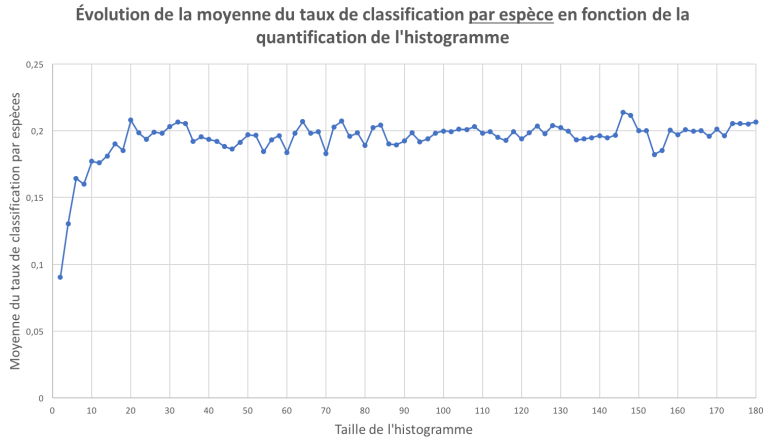


Figure 3.22 – Évolution du taux de classification moyen par espèce des images d'écorces classifiées avec le vecteur caractéristique complet en fonction de la taille de ce vecteur.

Les résultats de classification pour la base de 101 classes d'écorces avec un vecteur composé du vecteur de Gabor, de l'histogramme de teintes réduit à 30 bins, des mots verticaux et horizontaux sont présentés sur la Figure 3.23.

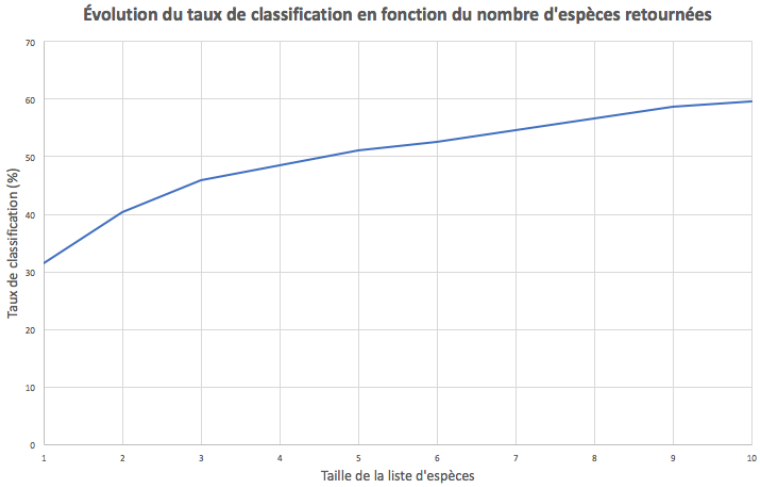


FIGURE 3.23 – Résultats de classification pour une base de 101 écorces d'arbres avec un histogramme de teintes réduit à 30 bins.

3.5 Comparaison

3.5.1 Comparaison sur la base de données *NewBarkTex*

La base de données d'écorces de 101 espèces correspond à la problématique de cette thèse, mais n'est pas une base de données de "référence", c'est-à-dire une base de données utilisée dans la littérature.

La base de données *NewBarkTex* [Por+14] est un jeu de données régulièrement utilisé lors du développement de descripteurs de textures notamment appliqués aux écorces [PVM07 ; Pal04 ; BYB15]. Cette base de données est constituée de 6 classes d'écorces possédant chacune 136 images de taille 64x64 (Figure 3.24).

(a) (b) (c) (d) (e) (f)

FIGURE 3.24 – Exemple d'images de la base *NewBarkTex*. a) *Betula pendula* b) *Fagus silvatica*, c) *Picea abies*, d) *Pinus silvestris*, e) *Quercus robus* et f) *Robinia pseudacacia*

Les particularités de cette base de données pour notre problématique sont le faible nombre de classes (contre près de 130 reconnues par *Folia*) et la faible variabilité au sein de chaque classe, ne montrant pas toute la diversité d'écorces au sein d'une espèce.

La Figure 3.25 montre le taux de classification obtenu sur cette base de données avec le vecteur de caractéristiques développé composé du vecteur de Gabor, de l'histogramme H, le mot vertical et le mot horizontal. Nous donnons les résultats pour le vecteur H complet (180 bins) et le vecteur réduit à 30 bins.

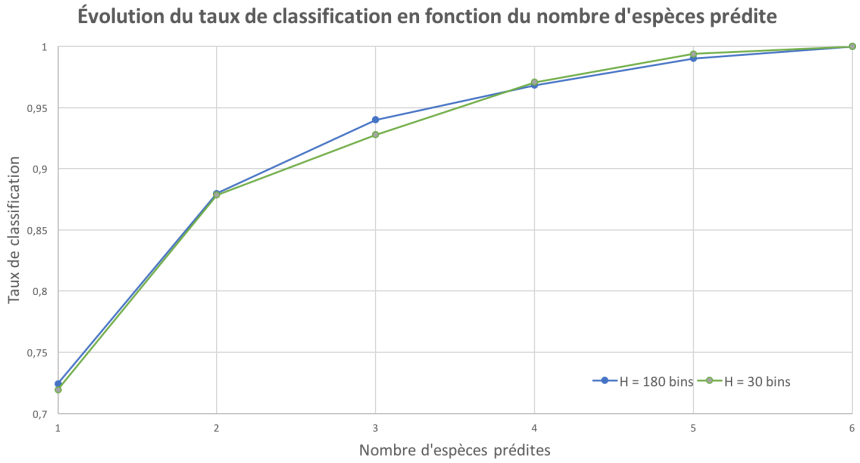


Figure 3.25 – Taux de classification sur la base de données NewBarkTex avec le vecteur complet pour deux valeurs de quantification de l'histogramme H.

Cette base de données est toujours utilisée par sa créatrice pour créer des vecteurs de caractéristiques de texture [Hoa+17]. Nous proposons donc de comparer notre vecteur d'écorces avec un vecteur basé LBP (Local Binary Patterns) utilisé dans la publication [Hoa+17]. L'information couleur d'un pixel est caractérisée par trois composantes de couleur dans un espace couleur tridimensionnel, dénoté $C_1C_2C_3$. L'opérateur LBP couleur consiste à attribuer à chaque pixel un label qui caractérise le motif local dans un voisinage. Chaque label est un nombre binaire calculé en seuillant la composante couleur des voisins en utilisant la composante couleur du pixel considéré. Le résultat du seuillage, effectué pour chaque pixel voisin, est ensuite codé par un masque de poids. De même que Ojala *et al.* [OPH96] lorsqu'ils introduisent l'opérateur LBP original, un voisinage de 3x3 pixels est pris en compte. Pour caractériser le motif local du pixel considéré, les valeurs pondérées sont finalement additionnées de sorte que chaque étiquette varie de 0 à 255. Afin de caractériser l'ensemble de l'image de texture couleur, le LBP est appliqué sur chaque pixel et pour chaque paire de composants. Les distributions correspondantes sont donc représentées dans neuf histogrammes différents : trois histogrammes LBP à l'intérieur d'un même composant ((C_1C_1), (C_2C_2) et (C_3C_3)), et 6 histogrammes entre les composants ((C_1C_2), (C_2C_1), (C_1C_3), (C_3C_1), (C_2C_3) et (C_3C_2)). Une texture de couleur est donc représentée dans un espace dimensionnel 9x256. Avec un tel vecteur appliqué dans l'espace colorimétrique HSV sur la base de données *NewBarkTex*, le taux de reconnaissance est de 70,47% [Hoa+17] (Table 3.1).

Table 3.1 – Taux de classification sur la base d'écorces *NewBarkTex* avec différents vecteurs de caractéristiques.

	LBP couleur [Hoa+17]	Notre vecteur	Notre vecteur avec H réduit à 30 bins
Taux de classification	70,47%	72,42%	71,94 %

Ainsi, le Tableau 3.1 montre que notre vecteur caractéristique d'écorces donne de meilleurs résultats de classification, tout en ayant une dimension plus faible (9x256 = 2304 bins pour le descripteur LBP contre 301 pour notre vecteur et 151 pour notre vecteur réduit).

3.5.2 Comparaison avec un descripteur de texture

La base de données *NewBarkTex* propose un échantillon très réduit d'écorces. Notre base de données de 101 classes d'écorces a été envoyée à Vinh Truong Hoang (co-auteur de [Hoa+17]) afin d'extraire le LPB de couleur précédemment comparé sur la base *NewBarkTex* [Hoa+17]. La classification est réalisée avec le classifieur SVM un contre tous (session 3.3.1). L'apprentissage est réalisé sur la même partie de la base de 101 espèces d'arbres et les tests sont réalisés sur la seconde partie. Les résultats de classification sont présentés sur la Figure 3.26.

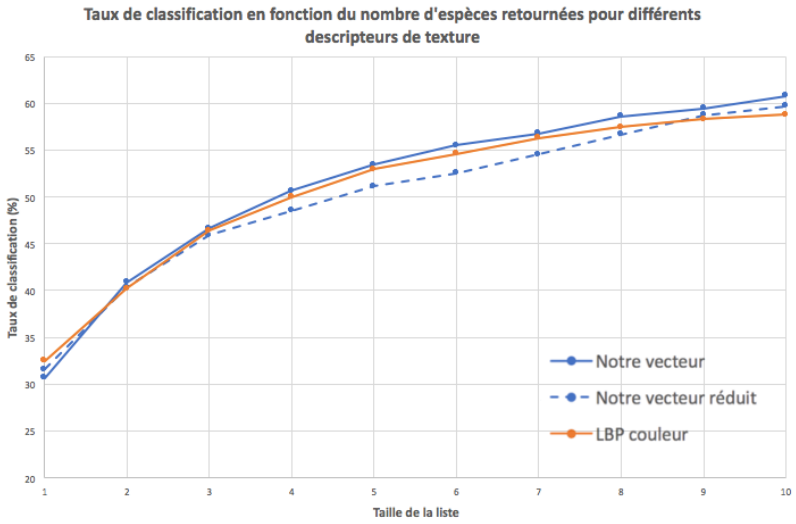


FIGURE 3.26 – Taux de classification sur la base de 101 classes d'écorces avec différents descripteurs de textures.

Nous pouvons voir que de meilleurs résultats sont obtenus avec le descripteur utilisant les LBP (courbe orange) dans le cas où une seule espèce est prédite. Mais le taux de classification obtenu avec le vecteur que nous avons développé donne des résultats assez proches (courbe bleue). Avec le vecteur complet (courbe bleue continue) les taux de classification sont supérieurs à ceux obtenues avec le vecteur de LBP lorsqu'une liste de plusieurs espèces est donnée. Ainsi, avec un vecteur de taille réduite (301), nous obtenons de meilleurs résultats de classification sur une base importante de 101 espèces d'arbres qu'avec un vecteur de dimension plus élevée (6912). Notre vecteur réduit (courbe bleue discontinue) rattrape les taux obtenus avec le LBP à partir d'une liste de 9 espèces prédites.

3.5.3 Comparaison avec le Deep Learning

Compte tenu du succès récent des réseaux de neurones convolutionnels, nous avons tenu dans cette thèse à comparer les taux de classification obtenus à l'aide de tels outils avec ceux que nous obtenons avec notre proposition. En effet, comme indiqué dans le Chapitre 2, les réseaux de neurones convolutionnels sont de plus en plus utilisés pour l'identification de végétaux. Bien que notre base de données possède peu d'échantillons, nous exposons dans cette partie les résultats de classification de deux réseaux de neurones largement utilisés.

Les CNN prennent en entrée des images de taille fixe. Les photos d'écorces dont nous disposons dans la base de données sont en mode portrait et ont des définitions différentes. Le risque de redimensionner ces images est d'avoir une déformation des structures de l'écorce. Pour pallier ce problème, des patchs de taille 100 par 100 sont découpés dans les images d'écorces avec un recouvrement de 50 pixels. En plus de mettre les données en forme pour les CNN, l'extraction de patchs permet d'augmenter le nombre de données. À partir de nos 1295 images d'apprentissage, 89324 patchs sont créés, dont 75% sont donnés en entrée des réseaux de neurones pour leur apprentissage, et 25% pour la validation. Sur la base de données de test, 88661 patchs sont créés et sont classifiés par les CNN afin d'évaluer leurs performances.

AlexNet

Le CNN AlexNet [KSH12] a une architecture assez simple de 5 couches de convolutions permettant l'extraction de descripteurs sur les images puis de 3 couches de neurones entièrement connectées pour la partie de classification (Figure 3.27). Ce réseau est entièrement entraîné sur notre base d'apprentissage.

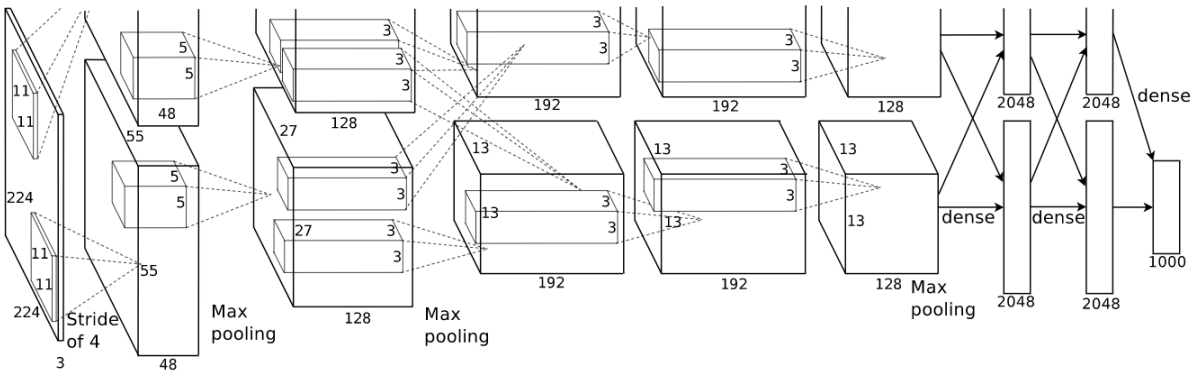


Figure 3.27 – Architecture du CNN AlexNet, illustration issue de [KSH12].

GoogLeNet

Le réseau de neurones à convolution GoogLeNet [Sze+15] a, depuis 2015, surpassé les performances du réseau AlexNet, grâce à une architecture beaucoup plus complexe (22 couches de convolution) (Figure 3.28). De même, ce réseau est entièrement entraîné sur notre base d'apprentissage.

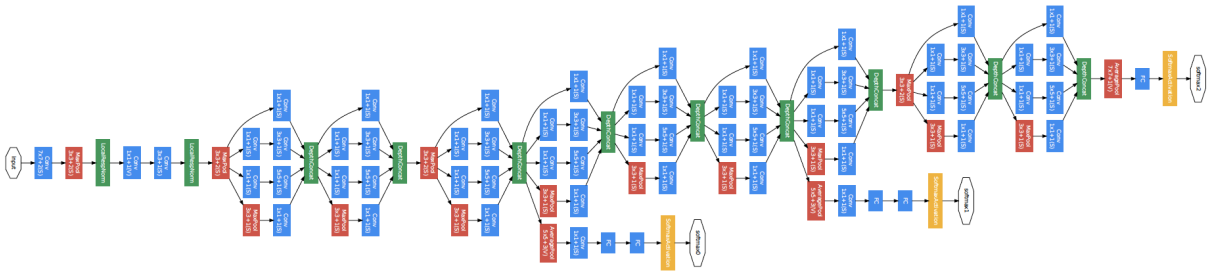


FIGURE 3.28 – Architecture du CNN GoogLeNet, illustration issue de [Sze+15].

Résultats de classification

Le Tableau 3.2 permet de comparer les taux de classification sur la même base de test obtenus avec les réseaux de neurones à convolutions AlexNet et GoogLeNet, et le vecteur que nous avons développé associé au classifieur SVN.

TABLE 3.2 – Comparaison des résultats de classification sur la base de 101 classes d'écorces avec des CNN.

	Top 1	Top 5	Moyenne par espèces
Notre vecteur	31,73%	52,48%	20,66%
AlexNet	41,46%	60,12%	32,08%
GoogLeNet	**44,21%**	**61,65%**	**33,36%**

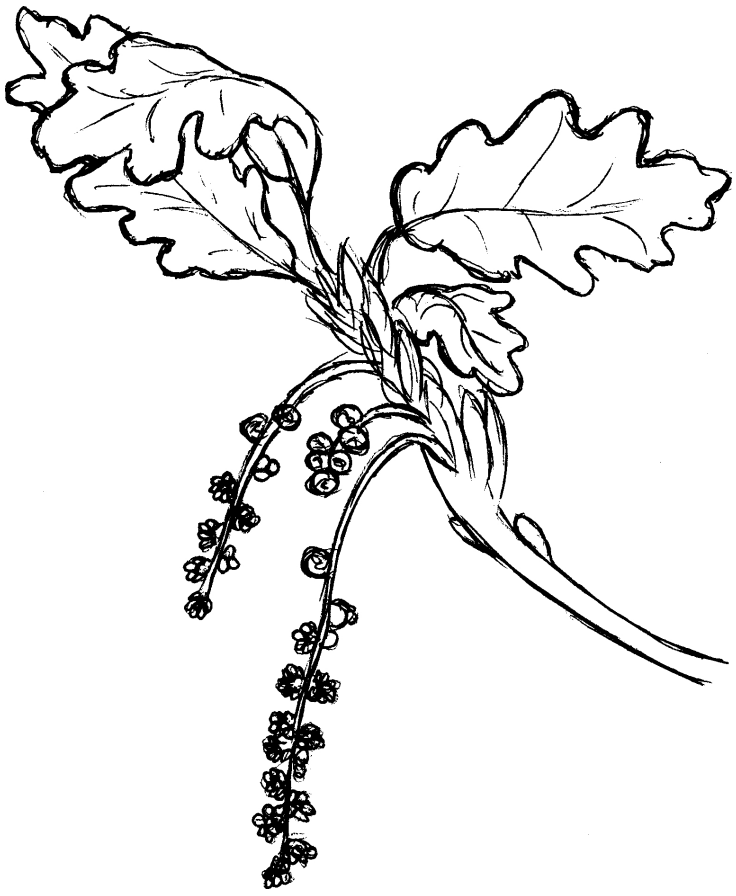
D'après ces taux de classification, nous pouvons voir que le vecteur que nous avons créé donne des résultats bien inférieurs aux réseaux de neurones à convolutions. Ceci confirme en outre la grande utilisation des CNN pour l'identification des végétaux [Jol+16b]. En comparant les résultats des deux CNN, nous pouvons déduire que le nombre de patchs extraits est suffisamment conséquent pour entraîner un réseau de neurones à convolutions important tel que GoogLeNet. Néanmoins, malgré des résultats largement supérieurs, les CNN ne fournissent aucune sémantique sur les éléments caractéristiques qu'ils extraient grâce à leurs nombreux filtres contrairement au vecteur créé "à la main". Nous reviendrons sur les réseaux convolutionnels au Chapitre 7 où nous esquisserons une méthode visant justement à extraire de la sémantique dans les CNN.

3.6 Conclusion

Dans ce chapitre, nous avons vu la création d'un vecteur caractéristique d'écorce qui se base sur les éléments morphologiques et colorimétriques utilisés par les botanistes pour identifier des espèces d'arbres. Les différents paramètres de ce vecteur ainsi que l'influence des différentes parties entre elles ont été présentés. Les résultats de classification sur deux bases de données d'écorces ont été exposés. La première est notre base de 101 classes, et la seconde est une base comportant peu de classes mais utilisée dans la littérature. La supériorité du vecteur d'écorce créé, en termes de performance et de dimensionnalité, a été démontrée sur ces deux bases de données par rapport à un descripteur de texture (LBP couleur). Cependant les performances de ce vecteur restent inférieures aux résultats de classification que nous pouvons obtenir avec des réseaux de neurones à convolutions. Cependant le principal avantage de notre vecteur caractéristique d'écorces est qu'il ne se comporte pas comme une "boîte noire", contrairement aux CNN.

Il faut cependant remarquer que les taux de classification des écorces sont relativement faibles, et qu'identifier une espèce d'arbre uniquement grâce à l'écorce est assez hasardeux. Dans la suite, nous souhaitons étudier l'identification des espèces d'arbres *via* les éléments botaniques caractéristiques des espèces que sont les fleurs, mais aussi les fruits.

Chapitre 4
Segmentation des fruits et fleurs



Ce chapitre décrit la méthode de segmentation qui a été développée dans l'objectif d'extraire une inflorescence ou un fruit du fond à partir d'une prise de vue en milieu naturel. Les fruits et les inflorescences ne sont pas traités indépendamment car, comme exposé dans le Chapitre 1 bis, un utilisateur novice ne peut indiquer si l'objet d'intérêt est une fleur ou un fruit. Ainsi, aucune connaissance *a priori* sur la nature de l'objet n'est considérée. Les différentes contraintes liées à l'objet d'étude sont exposées dans une première partie, avant de présenter en détail la méthode de segmentation.

4.1 Cadre de travail

Pour rappel, nous avons vu dans le chapitre traitant des éléments botaniques que l'identification végétale s'effectue en premier lieu à partir des organes reproducteurs des plantes, c'est-à-dire les fleurs. Les fruits étant l'évolution de la fleur fécondée, ceux-ci portent une certaine information. Il est donc très intéressant de considérer ces organes pour l'identification des espèces d'arbres. Dans notre cas, nous souhaitons extraire de l'information botanique à partir de ces fleurs et de ces fruits. Ainsi, une étape de segmentation est nécessaire afin de pouvoir extraire des caractéristiques comme la forme, la couleur ou la texture appartenant uniquement à l'objet d'intérêt.

La méthode de segmentation se doit d'être robuste aux différentes contraintes que posent la segmentation des fruits et des fleurs dans un environnement complexe. L'une des premières contraintes est l'utilisation du smartphone (faible résolution possible des images, angle de vue, *etc.*). L'algorithme de segmentation doit être en mesure de traiter des images prises en milieu naturel par des non professionnels dans des conditions d'éclairage non contrôlées. D'autre part, aucune information *a priori* ne peut être envisagée au vu de la multitude de formes, de couleurs et de textures des fruits et des fleurs (voir Figure 4.1). Ainsi, aucun modèle de forme déformable ne peut être appliqué comme cela peut être le cas quand seuls les feuilles [Cer+11], les fleurs [NZ07], ou les fruits [PJJ12] sont étudiés.

FIGURE 4.1 – Variabilité de forme, de couleur et de texture des fruits (les trois images de gauche) et des fleurs (les trois images de droite) des arbres présents en France métropolitaine.

Récemment, les smartphones ont commencé à intégrer deux caméras sur leur face arrière, ce qui augmente les possibilités d'acquisition et de traitement d'images, notamment

pour retrouver l'information de profondeur. Cette information de profondeur pourrait constituer une grande aide pour la segmentation car l'objet d'intérêt est très généralement au premier plan. Les systèmes de vision basés sur la stéréovision sont souvent utilisés en agriculture pour lutter contre les "mauvaises" herbes (espèces adventices) ou pour l'examen foliaire [Kaz+14], mais aussi pour la récolte des fleurs [Koh+11] ou des fruits [TZF02]. La principale contrainte du système de vision stéréoscopique est la nécessité d'une étape d'étalonnage. Étant donné la diversité de smartphones, cette étape ne peut pas être effectuée, d'autant plus que les paramètres des caméras sont inconnus. Cependant, avec des caméras non étalonnées, il est possible de mettre en place une méthode de co-segmentation. La co-segmentation consiste à segmenter un même objet dans un ensemble donné d'images. L'objet d'intérêt est photographié dans différents environnements [VRK11]. Chai *et al.* [CLZ11] utilisent une méthode de co-segmentation pour extraire les fleurs du fond des ensembles de données Oxford Flower 17 et Oxford Flower 102. La Figure 4.2 donne un exemple de co-segmentation réalisé sur le fruit du févier d'Amérique (*Gleditsia triacanthos*) avec l'algorithme de Dong *et al.* [Don+15]. Cependant, cette méthode n'est pas utilisée en agriculture pour l'extraction de feuilles, de fleurs ou de fruits car il faut au moins une image avec un arrière-plan différent.

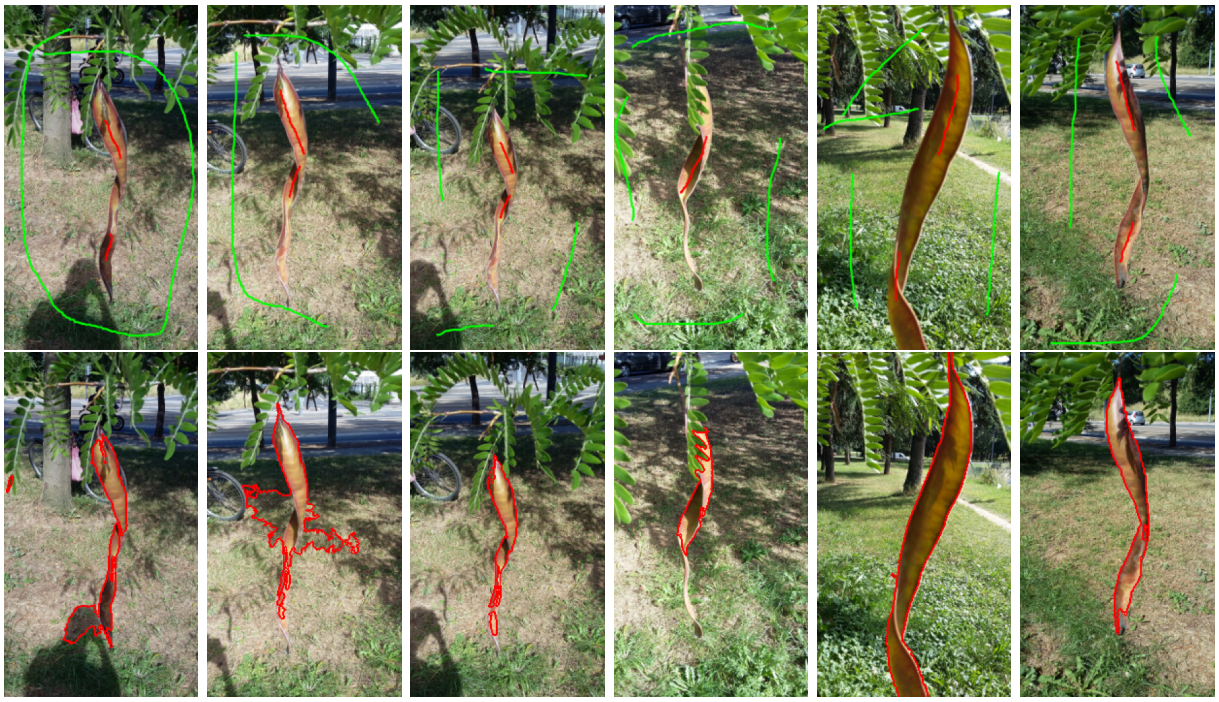


Figure 4.2 – Co-segmentation du fruit de *Gleditsia triacanthos*. La première ligne présente les images en entrée de l'algorithme de Dong *et al.* [Don+15] avec en vert les marques sur le fond et en rouge les marques sur l'objet. La seconde ligne expose les résultats de segmentation de l'objet pour chaque image d'entrée. C'est la cinquième image qui présente un arrière-plan différent.

La méthode proposée dans ce chapitre est basée sur la double acquisition d'image qui permet de récupérer des informations de profondeur mais sans avoir recours à la vision

stéréoscopique ou à la reconstruction 3D.

D'autre part, nous pouvons avoir accès à l'utilisateur qui, grâce à une application dédiée, peut avoir quelques interactions avec le processus, notamment en pointant l'objet d'intérêt, comme cela est le cas dans l'application *Folia*.

4.1.1 Création de la base de données

Au cours de cette thèse, nous avons utilisé un iPhone 7 Plus équipé de deux caméras en face arrière : une caméra de 12 MP (focale 28 mm, ouverture f/1.8) et une caméra 12 MP (focale 57 mm, ouverture f/2.8), respectivement intitulées "grand-angle" et "téléobjectif" par le constructeur Apple (voir Figure 4.3). Aucun paramètre des appareils photo, ni la distance entre eux ne sont connus par notre méthode, afin que celle-ci puisse fonctionner sur différents appareils.

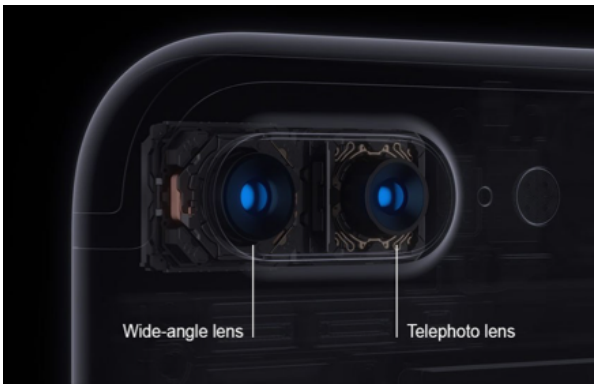


FIGURE 4.3 – Schéma des deux caméras en face arrière de l'iPhone 7 Plus.

Une application a été développée sur l'iPhone afin de prendre des photos brutes à partir des deux caméras. Il est important de noter que les deux images ne sont pas prises simultanément en raison d'une contrainte matérielle. En effet, à chaque fois que le contrôle d'une des deux caméras est demandé, la mise au point ainsi que la balance des blancs est effectuée. Il est donc clair que nous ne sommes pas dans une configuration stéréo-vision car les caméras ou l'objet d'intérêt peuvent se déplacer pendant cet intervalle. Pour réduire le temps entre les deux prises de vues, les étapes du processus d'acquisition sont les suivantes :

— Pointage de l'objet d'intérêt avec la caméra dite "téléobjectif" et mise au point automatique sur le centre de l'image. Les paramètres de mise au point et de balance des blancs sont sauvegardés.
— Un bouton (*AutoCAM2*) permet de passer sur la caméra dite "grand-angle" pour faire la mise au point sur celle-ci. Les paramètres de mise au point et de balance des blancs sont sauvegardés.
— Une fois la caméra "grand-angle" paramétrée, l'application repasse sur la caméra "téléobjectif". Un bouton (*Photos*) permet de déclencher les prises de vues sur les

deux caméras. L'intervalle est d'environ 200 ms.

— L'interface suivant la prise de photos permet de dessiner une trace utilisateur Λ sur l'image prise par le téléobjectif.

— Enfin, les deux photos ainsi que la trace utilisateur peuvent être validées avant sauvegarde.

Les différentes interfaces de l'application sont visibles sur la Figure 4.4.

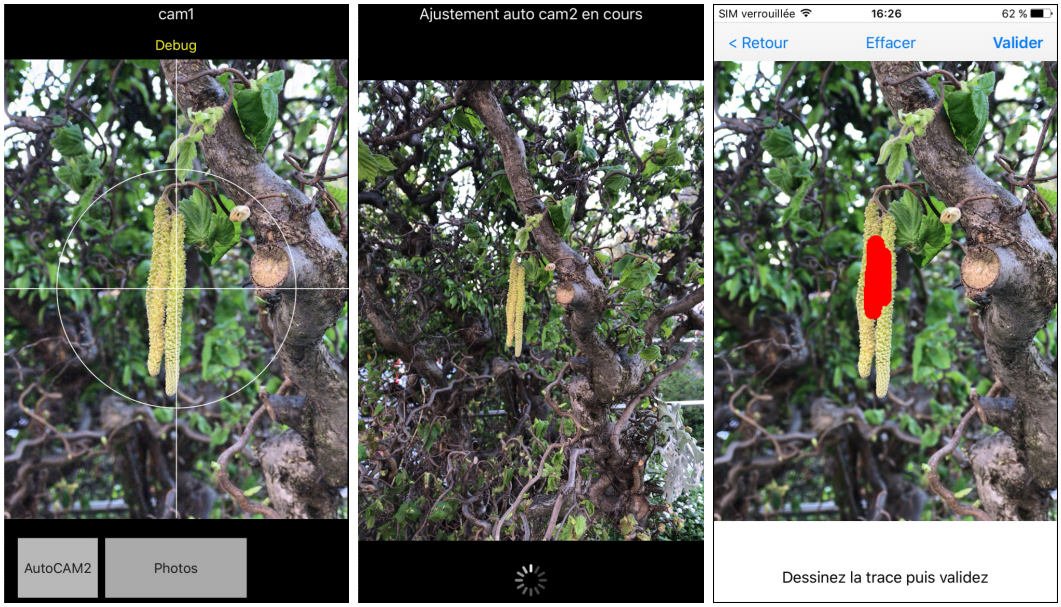
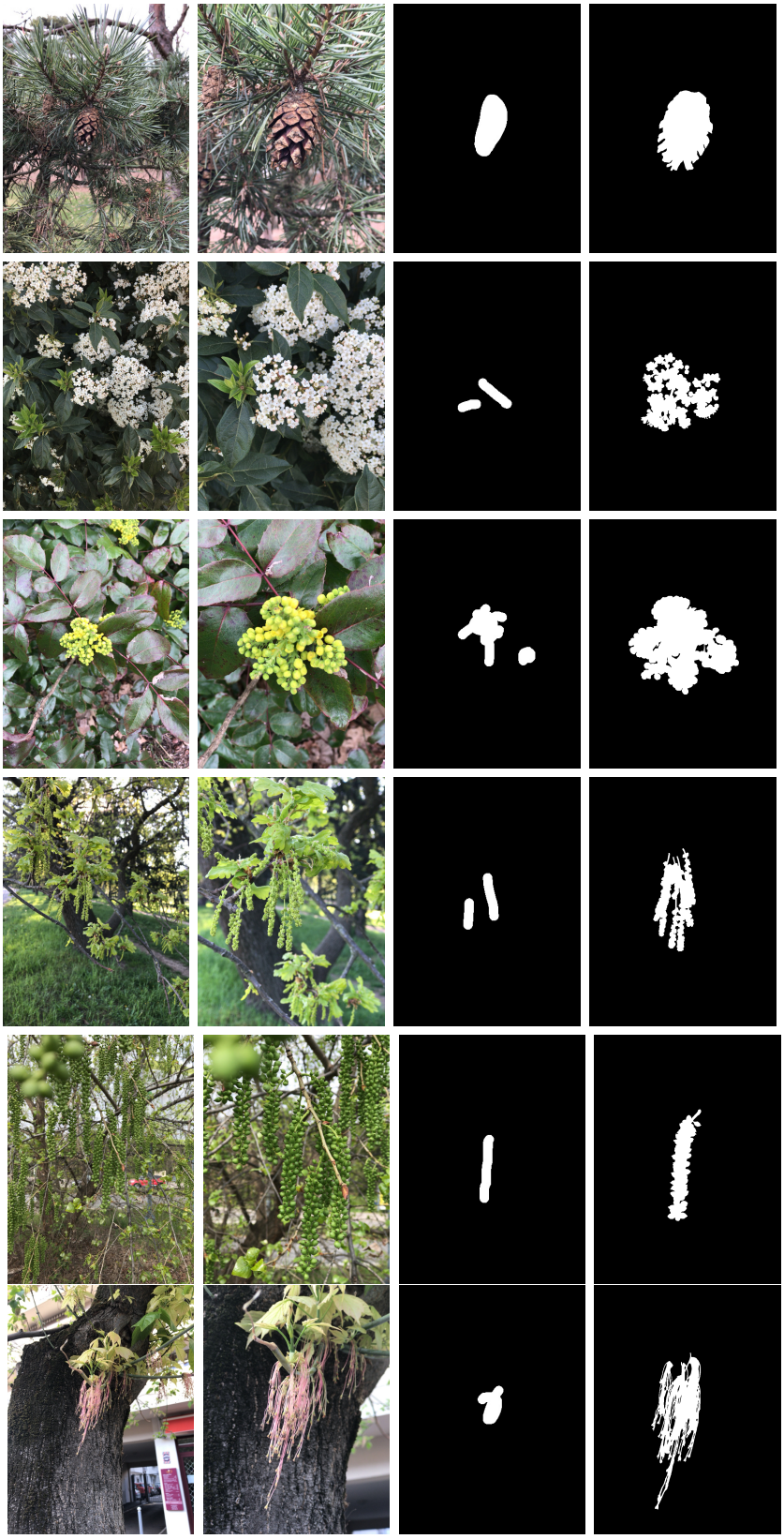


Figure 4.4 – Captures d'écran de l'application permettant de prendre en photo un organe avec chacune des deux caméras arrières de l'iPhone 7 Plus, ainsi que l'interface permettant de dessiner la trace utilisateur sur l'objet d'intérêt.

La base de données acquise est composée de 57 paires d'images de fruit ou d'inflorescence (de taille 768 x 1024 pixels) accompagnées d'une trace utilisateur pour chaque image "téléobjectif". Pour évaluer notre méthode de segmentation, les inflorescences et les fruits ont été segmentés manuellement sur chacune des 57 images "téléobjectif". Cette base de données a été rendue publique lors de la conférence ICMV 2018 (10^{th} *International Conference on Machine Vision*)[BCT18a]. La Figure 4.5 présente des exemples des paires d'images avec les traces utilisateurs, ainsi que le masque de l'organe segmenté. Il est à noter que cette base de données ne contient pas uniquement des inflorescences ou des fruits d'arbres feuillus mais aussi de conifères ou de plantes herbacées pour éviter le manque d'images.

Les temps de calcul donnés dans la suite de ce chapitre sont obtenus avec un MacBook Air possédant un processeur Intel Core i7 de 2,2 GHz et une mémoire DDR3 de 8 Go 1600 MHz.

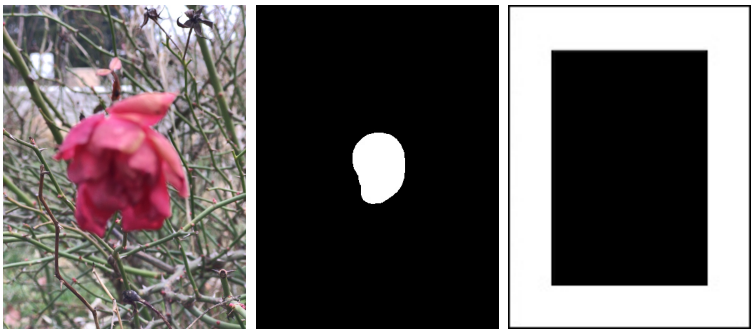


(a) (b) (c) (d)

FIGURE 4.5 – Exemples de photos prises avec a) la caméra "grand-angle", b) la caméra "téléobjectif", c) la trace utilisateur et d) l'image "téléobjectif" segmentée manuellement.

4.2 Carte de MLP

Bien que les fruits et les fleurs présentent une grande variété de couleur et de texture, ces caractéristiques diffèrent généralement du feuillage de la plante afin de faciliter la pollinisation et la récolte par les animaux. Dans un premier temps, nous cherchons donc à générer une carte de probabilité d'appartenance à l'objet d'intérêt, ou au fond, se basant sur des informations de texture ou de couleur. Grâce à la trace utilisateur, nous avons accès à un échantillon des pixels appartenant à l'objet d'intérêt. Bien que nous supposions que l'objet d'intérêt est au centre de la photo, il serait risqué de considérer uniquement les pixels du centre comme appartenant à l'objet. En effet, les fleurs et les fruits peuvent présenter plusieurs couleurs et textures que la trace utilisateur peut révéler. L'idée est donc de découper des patchs de taille $n * n$ dans l'image "téléobjectif" et de les étiqueter selon s'ils sont contenus dans le fond ou dans l'objet d'intérêt. Les patchs sont considérés comme appartenant à l'objet d'intérêt s'ils appartiennent à la trace utilisateur et comme appartenant au fond s'ils appartiennent au masque de fond (Figure 4.6). Le masque de fond correspond à 20% des bords de l'image.



(a) (b) (c)

Figure 4.6 – Image téléobjectif initiale a) avec la trace utilisateur b) et le masque de fond c).

Les patchs strictement inclus dans les masques sont conservés pour les données d'apprentissage qui sont séparées en deux classes : la classe objet et la classe fond. Sur la Figure 4.7, les patchs verts sont conservés et les patchs rouges rejetés. Le recouvrement des patchs est autorisé de manière à avoir un nombre d'imagettes suffisant pour l'entraînement du classifieur.

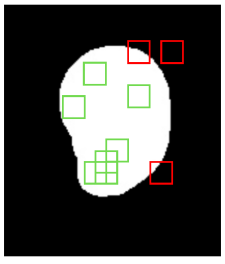


Figure 4.7 – Exemples de patchs conservés (verts) ou rejetés (rouges) pour la classe objet.

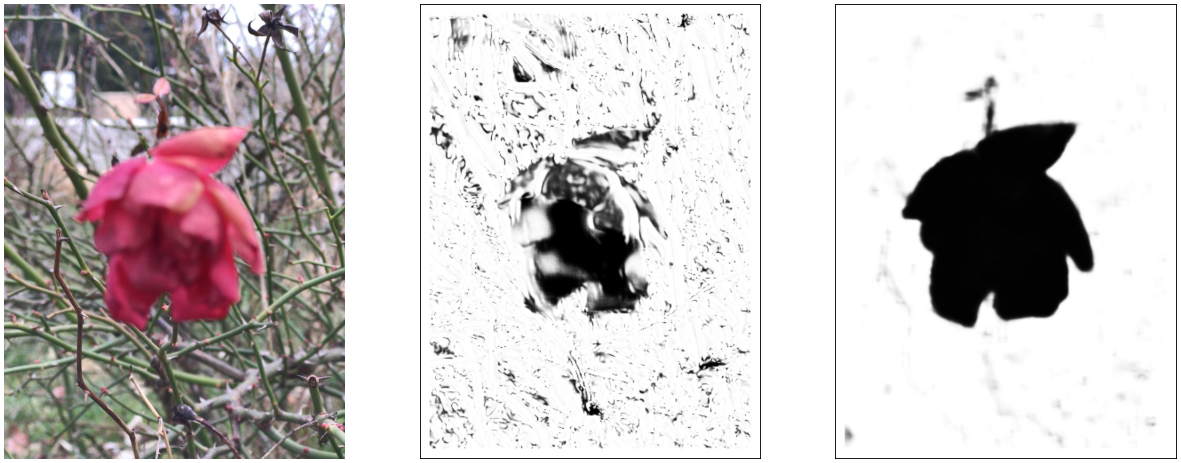
Ainsi, le classifieur est entraîné pour chaque image car les données d'apprentissage sont générées à chaque nouvelle acquisition. Afin de réduire le nombre de paramètres du classifieur, nous nous sommes orientés sur un classifieur MLP. Ici nous pouvons extraire un nombre suffisant de données (imagettes) sur l'image pour avoir un réseau de neurones performant. La contrainte principale est cependant d'avoir un réseau de faible complexité car embarqué sur le téléphone. Le nombre de neurones de la couche d'entrée correspond à la taille $n * n$ du patch multipliée par le nombre de canaux dans l'imagette. Le nombre de neurones sur la couche de sortie est de deux car nous avons deux classes : la classe de l'objet d'intérêt et la classe du fond. Concernant le nombre de neurones sur la seule couche cachée, une règle générale est d'avoir un nombre compris entre le nombre de neurones sur la couche d'entrée et le nombre de neurones sur la couche de sortie. Plus le nombre de neurones est élevé, plus il y a de paramètres à apprendre et donc un temps d'apprentissage plus long. Souhaitant un temps d'apprentissage le plus faible possible, ce nombre doit être faible. Dans un premier temps, le nombre de neurones sur la couche cachée est fixé à 5 de manière arbitraire.

Après la phase d'apprentissage, la carte de texture est calculée de la manière suivante : pour chaque pixel de l'image, un patch de taille $n*n$ est extrait autour du pixel. Ce patch est testé par le réseau de neurones précédemment entraîné qui va retourner pour chaque classe une probabilité d'appartenance entre 0 et 1. Ces valeurs peuvent être converties en niveau de gris pour donner une image. Nous appelons cette carte de disparité de texture D_{MLP}. En plus de la valeur de prédiction, le classifieur peut retourner pour chaque pixel de l'image si celui-ci appartient à la classe de fond ou à la classe de l'objet d'intérêt, constituant ainsi une image de segmentation binaire. Cette image est dénommée $D_{MLP\,binary}$. Dans un premier temps, afin d'intégrer assez de texture, nous prenons $n = 11$. Nous testons avec des patchs en niveau de gris et des patchs en couleur (espace colorimétrique RGB) afin d'étudier l'influence de la couleur dans les capacités de classification du réseau de neurones. Des premiers résultats sont visibles sur les Figures 4.8 et 4.9 .

Nous pouvons voir que l'information de couleur améliore, dans ces cas, significativement les résultats de classification entre la classe objet et la classe fond (Figure 4.9).

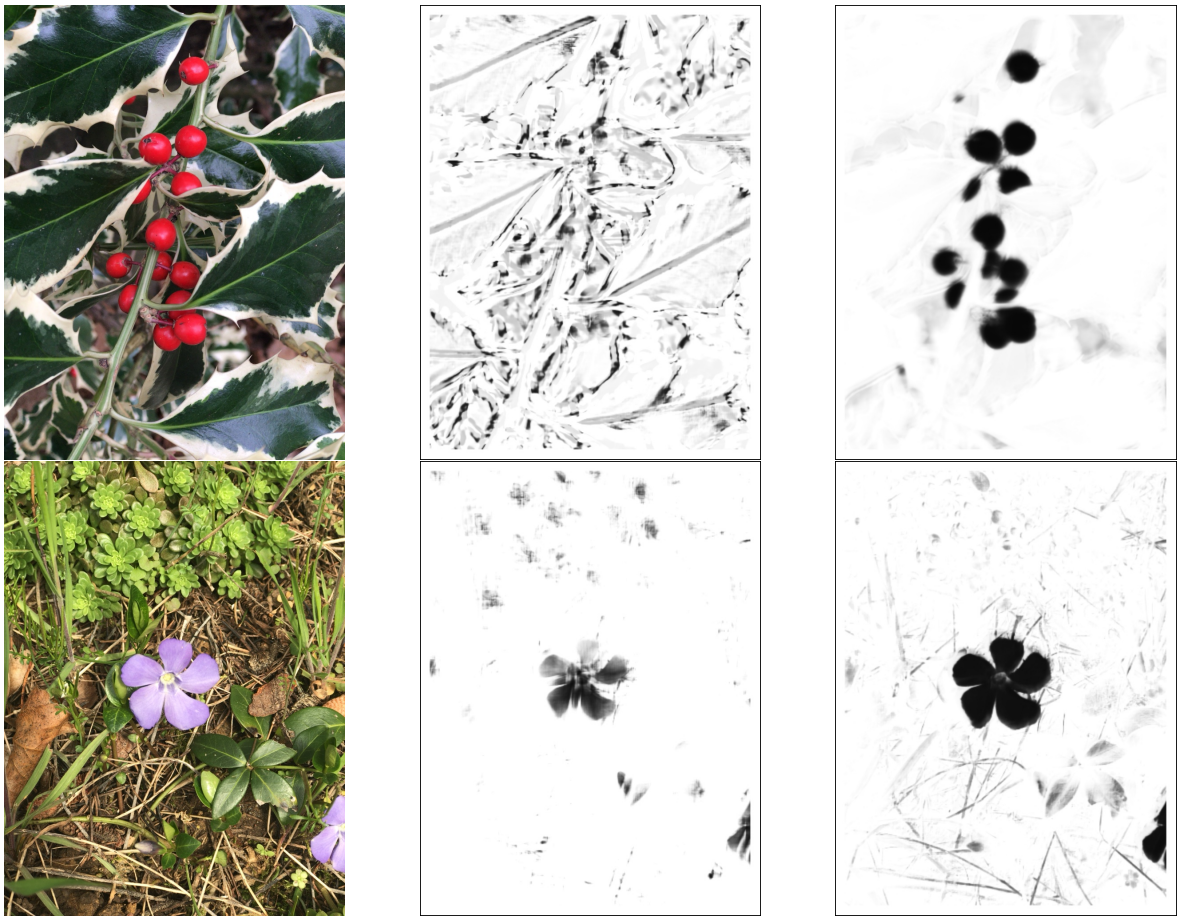
Pour une image de taille 1024x768, le temps pris par le programme pour créer les patchs est en moyenne de 0,42 seconde (patchs gris ou couleur). Le temps d'apprentissage du réseau de neurones est de 27 secondes si on prend l'image en niveau de gris, et 34 secondes pour une image couleur. Plus il y a de neurones dans le réseau, plus le temps d'apprentissage est long. Avoir une trentaine de secondes d'entraînement n'est pas souhaitable pour une application mobile. Pour réduire le temps d'apprentissage, le patch est réduit à une taille $n = 5$. Dans ce cas, le temps pris pour la construction des patchs est de 0,30 seconde et le temps d'apprentissage descend à 30 secondes pour une image en couleur et à 24 secondes pour une image en niveau de gris.

Sur la Figure 4.10, nous pouvons voir que la taille du patch n'influence pas beaucoup



(a) Image à segmenter. (b) D_{MLP} patchs niveaux de gris. (c) D_{MLP} patchs couleur.

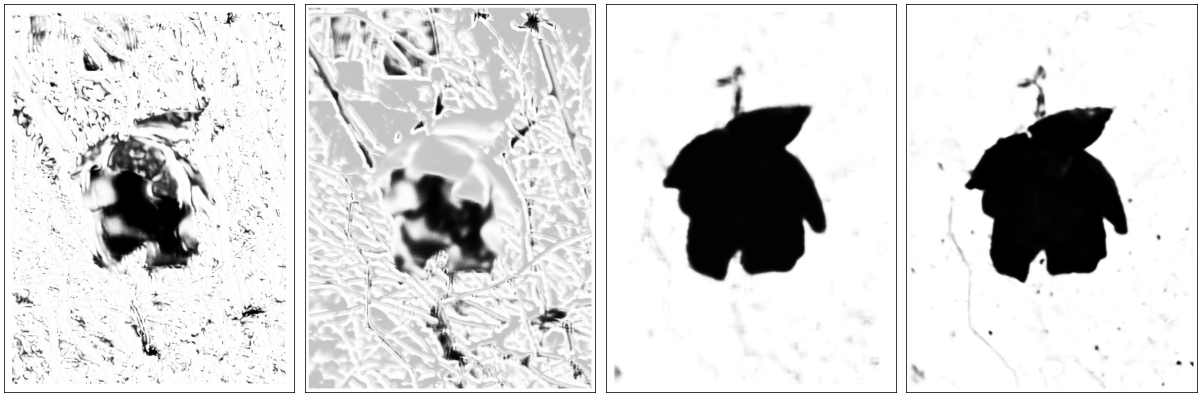
Figure 4.8 – Carte de texture de l'image initiale a) avec un patch pour $n = 11$ en niveau de gris b) et en couleur c).



(a) Image à segmenter. (b) D_{MLP} patchs niveaux de gris. (c) D_{MLP} patchs couleur.

Figure 4.9 – Autre exemples de carte de texture de l'image initiale a) avec un patch pour $n = 11$ en niveau de gris b) et en couleur c).

les résultats lorsque la couleur est conservée. Les contours sont cependant moins flous avec $n = 5$. Pour une image en niveau de gris, diminuer la taille du patch diminue les perfor-



(a) (b) (c) (d)

FIGURE 4.10 – Comparaison des cartes de texture entre des patchs en niveau de gris avec $n = 11$ a) et $n = 5$ b), et des patchs couleurs avec $n = 11$ c) et $n = 5$ d).

mances. Avoir une image en niveau de gris ne fait diminuer le temps d'apprentissage que d'environ 20% mais en échange d'une grosse perte de performance. Nous allons détailler d'autres pistes pour abaisser le temps d'entraînement du réseau de neurones.

4.2.1 Carte de couleurs

Dans le cas où l'objet d'intérêt (fleur ou fruit) est aussi représenté dans le fond, des patchs ressemblants à l'objet d'intérêt sont étiquetés comme appartenant à la classe fond, rajoutant alors de la confusion. Pour éliminer certains patchs du fond dont la couleur est proche de celle de l'objet d'intérêt, nous proposons d'utiliser une carte de dissimilitude des couleurs, notée D_C, exprimant pour chaque pixel la distance entre sa couleur et la couleur générale de l'objet. Le modèle de régression linéaire globale développé par Cerutti *et al.* [Cer+13b], actuellement utilisé dans l'application *Folia* pour caractériser la couleur de la feuille à segmenter, est utilisé pour calculer cette carte.

Les couleurs des pixels sont projetées dans l'espace colorimétrique L*a*b*. La droite de régression linéaire de cette distribution de couleur est calculée en prenant L (la luminosité) comme variable explicative. En prenant les couleurs des pixels contenus $(I_{Lab}(x, y))_{(x;y)\in\Lambda}$, avec Λ la trace utilisateur, nous pouvons obtenir les coefficients pour la prédiction de la couleur $\hat{I}_{Lab}(L, \Lambda) = L; \hat{a}(L, \Lambda); \hat{b}(L, \Lambda)$ à partir de deux fonctions affines :

$$\hat{a}(L, \Lambda) = a_0(\Lambda) + a_1(\Lambda)L \quad ; \quad \hat{b}(L, \Lambda) = b_0(\Lambda) + b_1(\Lambda)L \tag{4.1}$$

La similarité de couleur d'un pixel (x ;y) est la distance Euclidienne entre sa couleur $I_{Lab}(x; y)$ et la couleur prédite par la régression linéaire :

$$d_{LinReg,Lab}(x, y, \Lambda) = ((I_{Lab}(x, y) - \hat{I}_{Lab}(L(x, y), \Lambda))^T (I_{Lab}(x, y) - \hat{I}_{Lab}(L(x, y), \Lambda)))^{\frac{1}{2}} \tag{4.2}$$

Plus la distance est faible, plus la couleur du pixel est similaire aux couleurs contenues

dans la trace utilisateur. La Figure 4.11 présente des cartes de similarité de couleurs.

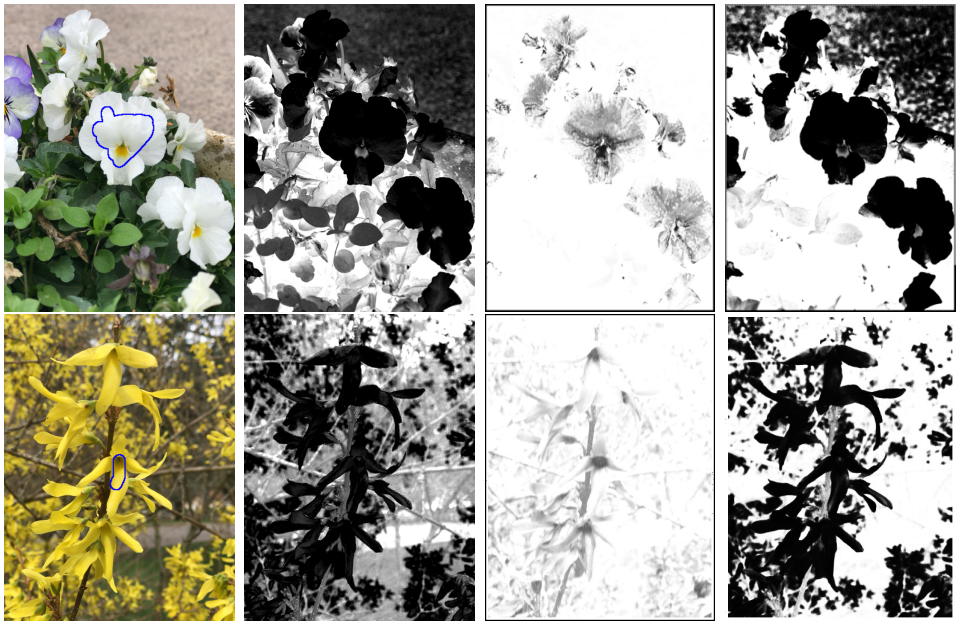
(a) (b) (c) (d) (e) (f)

Figure 4.11 – Exemples d'images initiales (a, c, e) avec le contour de la trace utilisateur en bleu, et les cartes de similarité de couleurs associées (b, d, f).

L'idée est d'utiliser cette carte de distance de couleur afin d'éliminer les patchs pour l'apprentissage appartenant à l'arrière-plan et ayant une couleur trop proche de l'objet d'intérêt. Plus précisément, les patchs qui, sur la carte de distance de couleur, ont un niveau de gris supérieur à la moyenne des niveaux de gris de la carte de couleurs dans la trace utilisateur sont éliminés. Supprimer ces patchs permet de diminuer le nombre de données d'apprentissage de la classe fond pour ainsi avoir des classes plus équilibrées, mais aussi de réduire le temps d'entraînement. Cette technique ne change pas les résultats quand il n'y a qu'un seul fruit ou une seule fleur dans l'image, mais améliore la distinction entre fruit/fleur et le fond lorsqu'il y en a plusieurs dans l'image, comme nous pouvons le voir sur la Figure 4.12.

En plus d'augmenter les performances du réseau de neurones, le temps d'apprentissage moyen diminue à 11 secondes (diminution de 63%). Le temps pour construire les données d'apprentissage est cependant rallongé de 0,09 seconde en moyenne à cause du calcul de la carte de couleurs, ce qui est malgré tout très rentable par rapport à la diminution du temps de calcul pour le réseau de neurones. D'autre part, nous pouvons observer que la carte de MLP donne des résultats plus contrastés que la carte de similarité de couleurs, justifiant ainsi son utilisation par rapport à la carte de similarité de couleurs seule (Figure 4.12 (b) et (d)).

Pour encore réduire la durée d'entraînement du réseau, la taille de l'image est divisée par deux. Cette durée passe donc à 4,3 s. Bien que les images soient réduites (taille 384 x 512), cela ne détériore pas les résultats du réseau de neurones. Le nombre de patchs appartenant à la classe fond donnés en entrée du réseau de neurones pour l'apprentissage est suffisamment élevé. On divise le nombre de patchs d'apprentissage de cette classe par quatre en limitant le recouvrement des patchs. Par contre, le nombre de patchs d'entraînement appartenant à l'objet d'intérêt reste le même. Jusqu'à maintenant, le nombre de neurones sur la couche cachée était fixé à 5. La réduction du nombre de neurones sur cette couche à 4 réduit le temps d'apprentissage sans variation visible des images de résultats. Les probabilités d'appartenance aux classes sont légèrement plus faibles, donc le réseau



(a) (b) (c) (d)

FIGURE 4.12 – Amélioration des cartes de MLP grâce à la prise en compte de la carte de similarité de couleurs. Image initiale avec la trace utilisateur en bleu a), carte de similarité de couleurs b), carte de MLP avant suppression des patchs du fond ayant une couleur trop similaire à l'objet d'intérêt c), et carte de MLP après suppression des patchs d).

de neurones répond avec moins de certitude. Si on réduit encore le nombre de neurones sur la couche cachée, on se met à ré-augmenter le temps d'apprentissage. Ainsi, la durée d'apprentissage est de 1,9 s. Avec la construction de la carte de MLP, le temps de calcul moyen pour obtenir une carte de MLP à partir d'une image "téléobjectif" et d'une trace utilisateur est de 2,87 secondes en moyenne sur notre base de données de 57 images. Des exemples sont présentés sur la Figure 4.13.

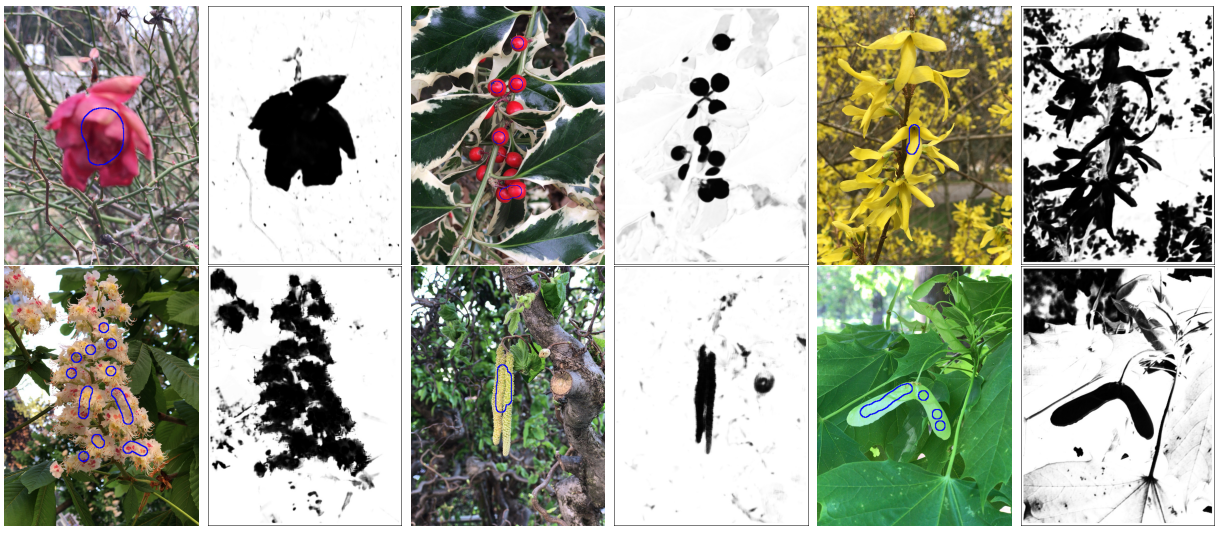


FIGURE 4.13 – Exemples de cartes de MLP obtenues à partir des images "téléobjectif" et des traces utilisateur (contour bleu).

Pour résumer, le schéma sur la Figure 4.14 représente le processus proposé pour obtenir la carte de similitude.

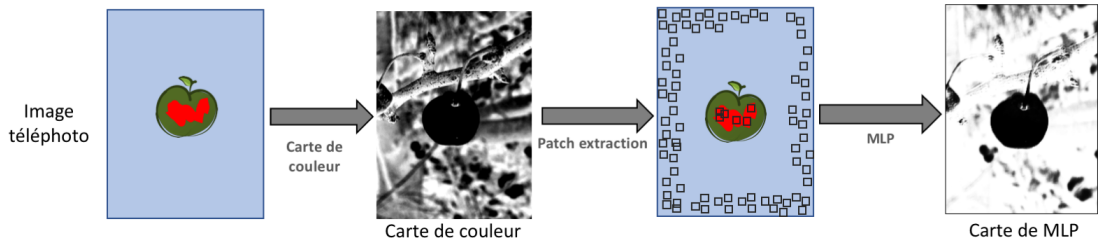
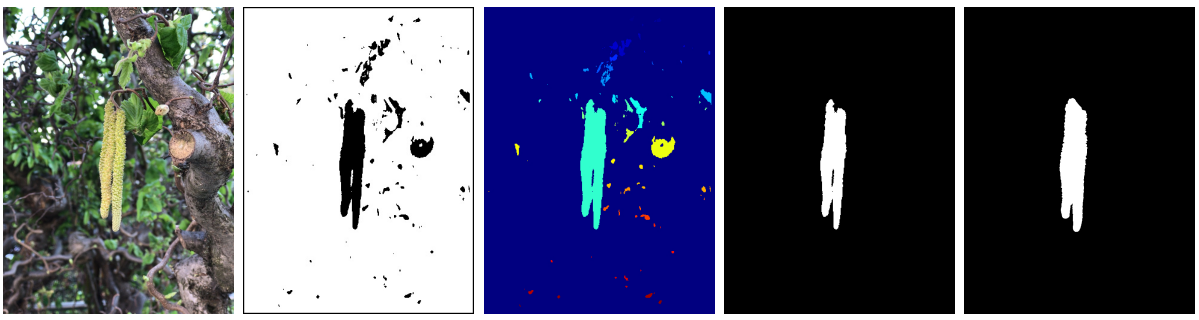


Figure 4.14 – Schéma de la génération de la carte de MLP.

4.2.2 Performance de la carte de MLP

Nous pouvons voir dans un premier temps que les cartes de MLP donnent de bons éléments pour effectuer la segmentation de fruits ou d'inflorescence dans une image. Comme énoncé précédemment, nous pouvons obtenir du réseau de neurones une image binaire $D_{MLP\,binary}$ qui indique, pour chaque pixel, si celui-ci appartient à l'objet d'intérêt ou au fond. C'est ici cette segmentation qui va être évaluée.

Il existe dans $D_{MLP\,binary}$ plusieurs composantes connexes correspondant à des objets détectés comme similaires à l'objet d'intérêt. Pour extraire l'objet d'intérêt, seules les composantes connexes intersectant la trace utilisateur sont conservées. Nous appelons cette méthode CC_{MLP} dans la suite.



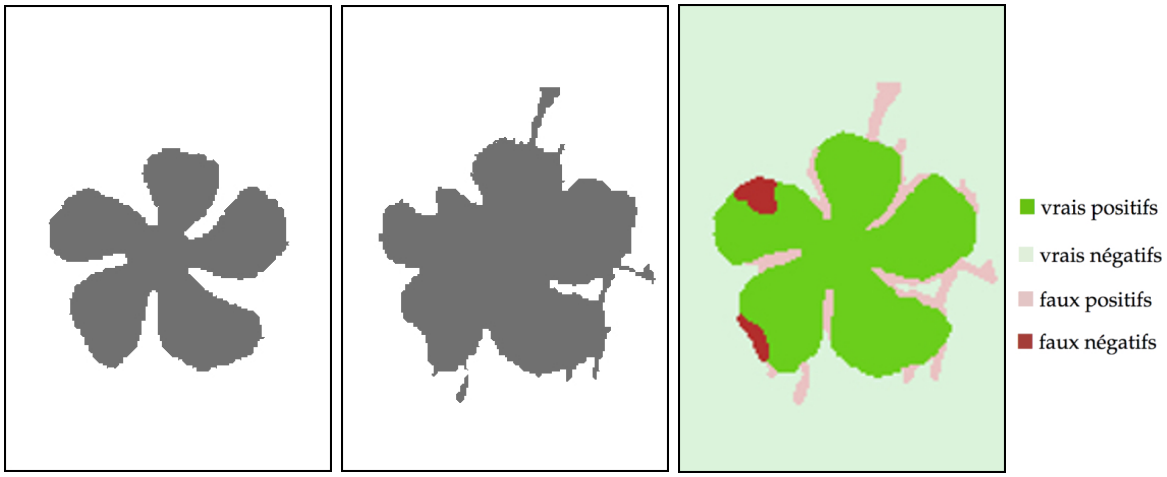
(a) (b) (c) (d) (e)

Figure 4.15 – Illustration de CC_{MLP} avec l'image "téléobjectif" en a), l'image binaire en sortie du MLP $D_{MLP\,binary}$ b), les composantes connexes c), la composante connexe intersectant la trace utilisateur d) et la segmentation manuelle e).

Grâce aux images manuellement segmentées, nous pouvons juger la qualité de la segmentation obtenue avec la méthode CC_{MLP} *via* les mesures de qualité que sont la précision, le rappel, l'indice de Dice, l'indice de Jaccard, et l'indice de Manhattan.

Nous rappelons ces mesures statistiques utilisées pour comparer deux ensembles. Nous notons Ω_{obj} la région de l'objet d'intérêt et Ω_{seg} la région segmentée par la méthode. La Figure 4.16 offre une représentation visuelle de ces différentes mesures.

— Vrais positifs (VP) : les pixels de Ω_{seg} présents dans Ω_{obj} : $VP = \Omega_{seg} \cap \Omega_{obj}$.
— Faux positifs (FP) : les pixels de Ω_{seg} absents de Ω_{obj} : $FP = \Omega_{seg} \cap \overline{\Omega_{obj}}$.
— Faux négatifs (FN) : les pixels absents de Ω_{seg} mais appartenant à Ω_{obj} : $FN = \overline{\Omega_{seg}} \cap \Omega_{obj}$.
— Vrais négatifs (VN) : les pixels absents de Ω_{seg} et de Ω_{obj} : $VN = \overline{\Omega_{seg}} \cap \overline{\Omega_{obj}}$.



(a) (b) (c)

FIGURE 4.16 – Visualisation des indices c) pour comparer une région segmentée Ω_{seg} b) par rapport à la vérité terrain Ω_{obj} a). Illustration issue de [Cer13a]

Les mesures de qualité citées précédemment utilisent les indices de bases que sont les vrais positifs, faux positifs, faux négatifs et vrais négatifs. Dans tous les cas, si la segmentation est parfaite, c'est-à-dire que $\Omega_{seg} = \Omega_{obj}$, la valeur des mesures est égale à 1. Ces mesures sont alors calculées comme suit :

— La précision est le rapport entre le nombre de vrais positifs et le nombre de vrais et faux positifs.

$$Precision(\Omega_{seg}, \Omega_{obj}) = \frac{VP}{VP + FP} \tag{4.3}$$

— Le rappel (ou sensibilité) est le rapport entre le nombre de vrais positifs et le nombre de pixels à retrouver.

$$Rappel(\Omega_{seg}, \Omega_{obj}) = \frac{VP}{VP + FN} \tag{4.4}$$

— L'indice de Dice (ou F-mesure) est le rapport entre deux fois le nombre de vrais positifs, et la somme du nombre de positifs et de pixels à trouver.

$$Dice(\Omega_{seg}, \Omega_{obj}) = \frac{2VP}{2VP + FP + FN} \tag{4.5}$$

— L'indice de Manhattan est le rapport entre le nombre pixels corrects, et le nombre

total de pixels.

$$Manhattan(\Omega_{seg}, \Omega_{obj}) = \frac{VP + VN}{VP + VN + FP + FN} \tag{4.6}$$

— L'indice de Jaccard (ou recouvrement) est le rapport entre le nombre de vrais positifs et la réunion de pixels positifs et à retrouver.

$$Jaccard(\Omega_{seg}, \Omega_{obj}) = \frac{VP}{VP + FP + FN} \tag{4.7}$$

Chacune de ces métriques permet de caractériser la qualité de la segmentation. Notamment, le rappel permet de valider qu'au moins tous les pixels de la vérité terrain ont bien été segmentés, mais ne considère pas les pixels segmentés alors qu'ils n'appartiennent pas à la vérité terrain (les faux positifs). Ainsi, le rappel va être élevé (proche de 1) dans le cas d'une sur-segmentation de l'objet d'intérêt. C'est la précision qui va pouvoir mettre en évidence une sur-segmentation par rapport à la valeur du rappel (la valeur de la précision sera faible). Les indices de Dice, de Manhattan et de Jaccard permettent de mesurer la similarité entre les deux régions de façon plus globale, puisque seul l'identité parfaite entre les deux régions donne une mesure égale à 1.

Les valeurs moyennes des différentes métriques pour la méthode de segmentation par composantes connexes de la carte binaire de MLP, notée CC_{MLP}, obtenue sur l'image téléobjectif, sont présentées dans le Tableau 4.1 pour l'ensemble de la base.

Table 4.1 – Mesure de qualité de la segmentation avec la méthode CC_{MLP}.

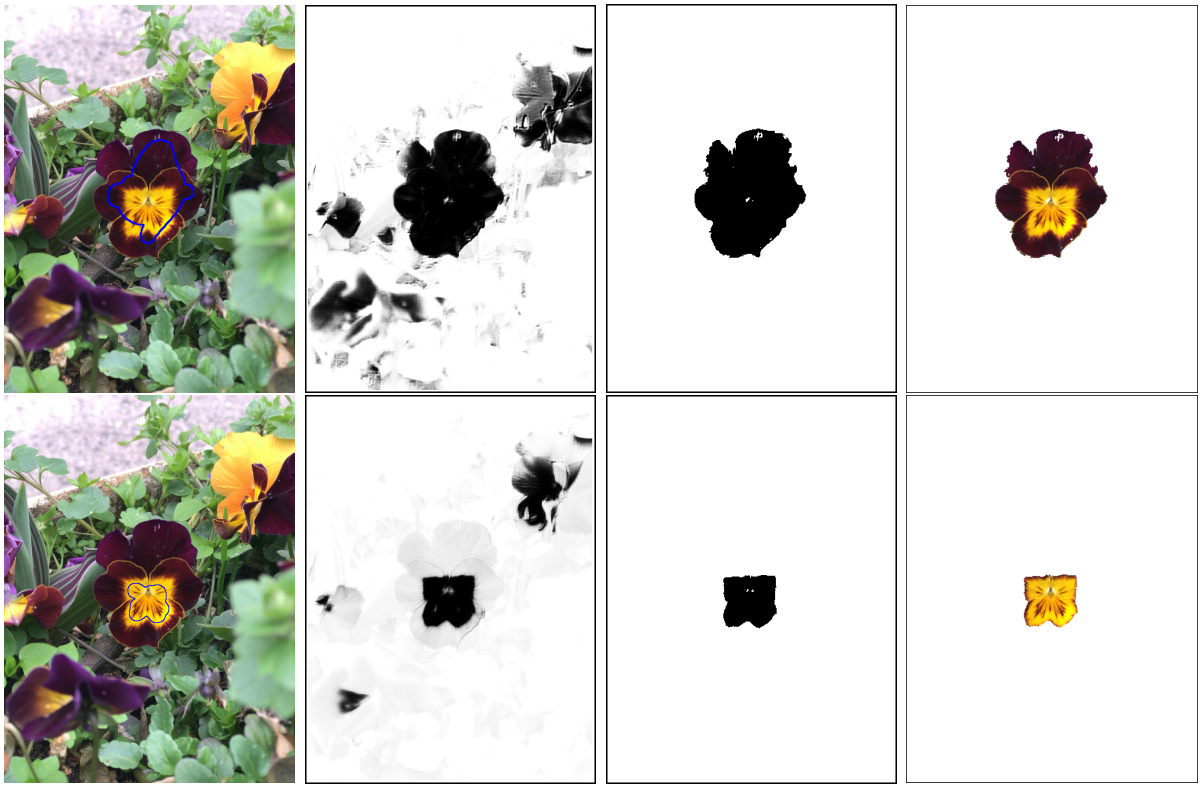
	Précision	**Rappel**	**Dice**	**Manhattan**	**Jaccard**
CC_{MLP}	0.6519	0.9362	0.7192	0.9439	0.6163

Bien que cette méthode donne de bons résultats visuels comme le montre la Figure 4.15, les mesures de précision et de rappel montrent que l'objet d'intérêt est en moyenne sur-segmenté, par exemple quand des objets similaires sont proches du fruit ou de l'inflorescence à segmenter. Nous allons voir dans la partie 4.4 comment palier ces éventualités.

4.2.3 Influence de la trace utilisateur

Au préalable, cette partie présente l'influence de la trace utilisateur sur la carte de MLP. Une grande importance est donnée à celle-ci car elle permet d'étiqueter les pixels appartenant à l'objet d'intérêt. Si la trace utilisateur ne recouvre pas toutes les couleurs de l'inflorescence ou du fruit, il manquera des informations au MLP pour effectuer un bon apprentissage. La Figure 4.17 donne un exemple de segmentation de la même fleur avec deux traces utilisateur différentes.

Nous ne traitons pas de l'influence de la trace utilisateur sur la carte de couleurs (se reporter à la thèse de Guillaume Cerutti [Cer13a]). Nous allons étudier l'évolution des



(a) (b) (c) (d)

FIGURE 4.17 – Résultats de segmentation d) différents en fonction de la trace utilisateur en entrée a) avec la carte de MLP b) et CC_{MLP} c).

mesures de qualité de segmentation utilisées précédemment en fonction d'une diminution ou d'une augmentation de la taille de la trace utilisateur. Pour cela, le masque correspondant à la trace utilisateur a été érodé ou dilaté. Un élément structurant de forme sphérique avec un diamètre de trois pixels est utilisé pour effectuer ces deux opérations morphologiques. La trace utilisateur est érodée jusqu'à ce qu'elle soit réduite de $x\%$ par rapport à la trace initiale, ou est dilatée jusqu'à ce que celle-ci soit étendue de $x\%$ par rapport à la trace initiale. Les résultats sont présentés sur les Figures 4.18 et 4.19.

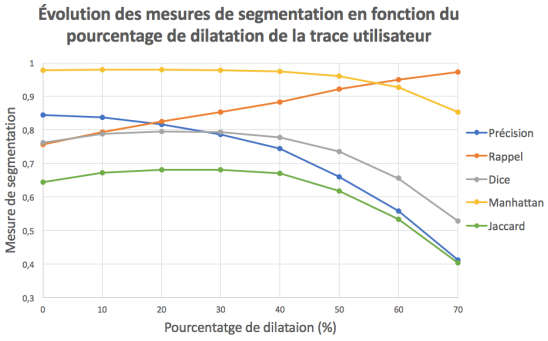


FIGURE 4.18 – Influence de la dilatation de la trace utilisateur sur la méthode CC_{MLP}.

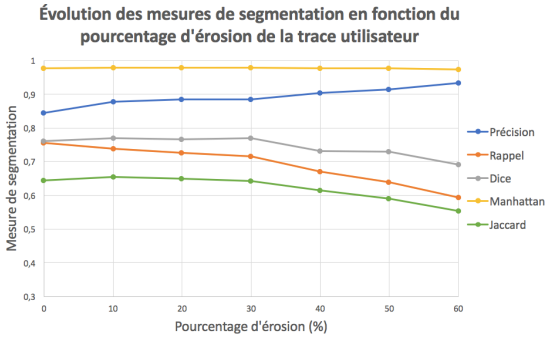


FIGURE 4.19 – Influence de l'érosion de la trace utilisateur sur la méthode CC_{MLP}.

La dilatation représente un utilisateur qui aurait dépassé de l'objet d'intérêt tandis que l'érosion correspond à un utilisateur qui n'aurait pas laissé une marque "raisonnable" sur l'objet d'intérêt. Le graphe de la Figure 4.18 nous indique que la dilatation de la trace utilisateur entraîne une sur-segmentation de l'objet d'intérêt. En effet, la précision diminue alors que le rappel augmente. Les évolutions de ces deux mesures de qualité indiquent que le nombre de faux positifs augmente et que le nombre de faux négatifs diminue. Les indices de Dice, Manhattan et Jaccard restent dans un premier temps constants car le nombre de faux positifs augmente et le nombre de faux négatifs diminue. Cependant, à partir d'une dilatation de 40%, la trace utilisateur commence à trop dépasser de l'objet d'intérêt pour intégrer des éléments du fond, ne permettant pas au MLP de distinguer correctement les deux classes. D'autre part, avec une trace utilisateur plus étendue, le nombre de patchs correspondant à l'objet d'intérêt pouvant être extrait est augmenté. Ceci entraîne une augmentation du temps de calcul.

Concernant l'érosion de la trace utilisateur, les évolutions de la précision et du rappel montrent une sous-segmentation de l'objet d'intérêt, à cause d'une diminution du nombre de vrais positifs et une augmentation des faux positifs. Éroder la trace utilisateur diminue le nombre de données d'apprentissage pour la classe de l'objet d'intérêt. Il en résulte, pour certaines fleurs ou fruits de petite taille avec initialement un faible nombre de patchs extraits, très peu de pixels classifiés comme appartenant à l'objet d'intérêt (nombre de vrais positifs proche de zéro). Sur notre base de données, le nombre de patchs extraits de la trace utilisateur sans dilatation ou érosion représente en moyenne 22% de tous les patchs d'apprentissage.

Pour conclure, les résultats de segmentation sont très dépendants de la trace utilisateur. Dilater celle-ci risque de sur-segmenter l'objet alors que l'éroder risque de sous-segmenter l'objet d'intérêt. Les bons résultats de la carte de MLP dépendent donc grandement de l'application de l'utilisateur à dessiner une trace recouvrant toutes les couleurs de l'objet d'intérêt sans trop dépasser sur le fond.

4.3 Carte de profondeurs

Dans un second temps, nous tirons parti de la présence des deux caméras afin de calculer une estimation de ce que pourrait être une carte de profondeurs. Comme le montrent les Figures 4.4 et 4.5, les deux photos ne sont pas dans le même espace géométrique à cause des jeux de lentilles qui sont différents pour les deux caméras. Afin de pouvoir accéder à l'information de profondeur, il faut que ces deux images soient dans le même espace géométrique de manière à pouvoir simuler deux prises de vues avec un faible décalage. Par conséquent, comme étape préliminaire, l'image "grand-angle" est recalée sur l'image "téléobjectif". Le recalage d'image consiste à mettre en correspondance deux images en vue de leur combinaison. Cette combinaison s'effectue par la recherche de la transformation géométrique permettant de passer de l'image "grand-angle" à l'image "téléobjectif".

Ce sont les points d'intérêt et les descripteurs SURF (Speeded Up Robust Features) [BTV06] qui ont été utilisés pour faire la correspondance entre les deux images. SURF a été privilégié par rapport à SIFT notamment car celui-ci est plus rapide en temps de calcul. Plus précisément, nous utilisons la version upright-SURF (U-SURF) qui ne tient pas compte de l'orientation. En effet, il n'y a pas de rotation des caméras pendant la prise de vue. Ne pas tenir compte de l'orientation permet de réduire les temps de calculs. Le détecteur SURF ne fonctionne que sur des images avec un seul canal, donc les deux images sont transformées en niveaux de gris. Nous souhaitons faire le recalage sur l'objet d'intérêt. Il est donc inutile d'extraire des points d'intérêt sur l'ensemble de l'image. Cela augmenterait les temps de calcul tout en diminuant les résultats. Ainsi, seuls les points caractéristiques extraits dans le centre des images (dans un cercle de rayon égal à 1/4 de la plus faible dimension de l'image) sont considérés. Pour détecter si un contour est conservé comme un point d'intérêt, le déterminant de la matrice Hessienne est calculé. Si ce déterminant σ est supérieur à un seuil, alors le point d'intérêt est conservé. Ce seuil est dans un premier temps assez élevé de manière à ce que les points caractéristiques détectés soient bien discriminants et sur les contours de l'objet. Des exemples de points extraits sur une paire d'images sont disponibles sur la Figure 4.20.

(a) (b) (c)

FIGURE 4.20 – Extraction des points d'intérêts SURF sur l'image grand-angle en haut et l'image téléobjectif en bas a) sur l'ensemble de l'image b) et restreint au centre c).

Une fois que les points caractéristiques et les descripteurs SURF sont extraits de l'objet des deux images, un appariement doit être fait. L'appariement Brute-Force est simple : il prend le descripteur d'un point caractéristique du premier ensemble et l'apparie avec toutes les autres points caractéristiques du deuxième jeu en utilisant un calcul de distance. La bibliothèque FLANN (Fast Library for Approximate Nearest Neighbors) [ML09] contient une collection d'algorithmes optimisés pour la recherche rapide du voisin le plus proche dans les grands ensembles de données et fonctionnant plus rapidement que le Brute-Force dans la plupart des cas. L'appariement FLANN que nous avons utilisé construit une structure de données efficace (KD-Tree) qui sera utilisée pour rechercher un voisin approximatif, tandis que l'appariement Brute-Force fait une recherche exhaustive et garantit de trouver le meilleur voisin. Seuls les 30 meilleurs appariements sont conservés, c'est-à-dire les paires de points caractéristiques dont la distance euclidienne entre les descripteurs est la plus faible (Figure 4.21).

Figure 4.21 – Mise en correspondance des 30 paires de points caractéristiques présentant les meilleurs appariements.

À partir de ces 30 paires de points, nous pouvons construire la matrice d'homographie. Une homographie est une transformation qui met en correspondance les points d'une image avec les points correspondants de l'autre image. Deux images sont liées par une homographie si les deux images montrent la même scène mais avec un angle différent. Dans notre cas, les angles de prises de vues sont bien différents car les deux caméras sont

côte à côte. La matrice d'homographie s'écrit de la manière suivant :

$$H = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \tag{4.8}$$

Considérons le premier ensemble de points correspondants (x_1, y_1) dans la première image et (x_2, y_2) dans la deuxième image. Ensuite, l'homographie H les apparie ensemble de la manière suivante :

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = H \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{4.9}$$

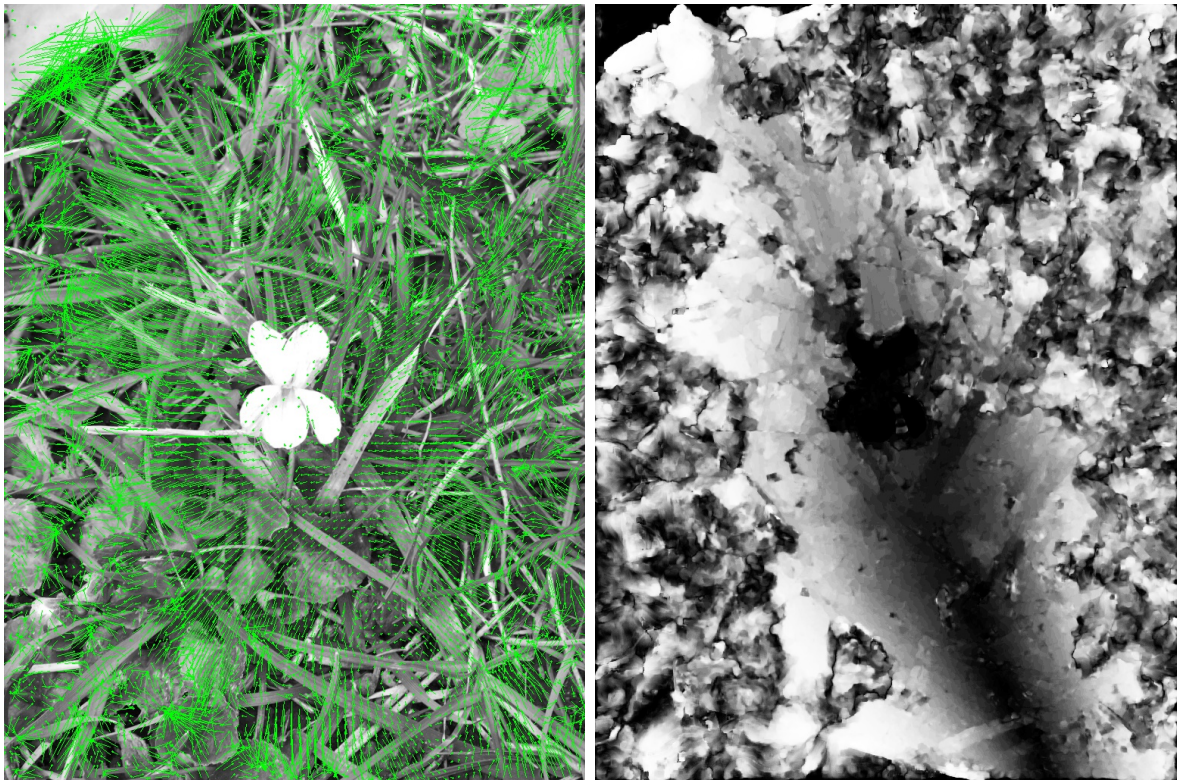
Par la suite, cette matrice est utilisée pour appliquer une transformation de perspective à l'image grand-angle (Figure 4.22).

(a) Image grand-angle (b) Image téléobjectif (c) Transformation homographique

FIGURE 4.22 – Visualisation c) de l'application sur l'image grand-angle de la matrice d'homographie calculée entre l'image grand-angle a) et l'image téléobjectif b).

Une fois que l'image grand-angle est dans le même espace géométrique que l'image téléobjectif, nous nous retrouvons avec deux images stéréo artificielles. À partir de ces deux images, le flux optique est calculé. Le flux optique est le mouvement apparent des objets de l'image entre deux images consécutives causé par le mouvement de l'objet ou de la caméra. Il s'agit d'un champ vectoriel 2D où chaque vecteur est un vecteur de déplacement montrant le mouvement des points de la première image à la seconde. Nous souhaitons calculer le déplacement pour tous les pixels. Pour cela, l'algorithme de Farneback a été utilisé [Far03]. Cet algorithme calcule le déplacement des pixels selon la direction verticale (Δ_y) et la direction horizontale (Δ_x). La Figure 4.23a montre le flux optique des pixels entre l'image téléobjectif et l'image grand-angle transformée dans l'espace géométrique de l'image téléobjectif. Pour faciliter la visualisation, le vecteur de flux optique est dessiné

pour chaque pixel avec un pas de dix pixels. Les zones où il n'y a pas de vecteurs dessinés correspondent à un flux optique nul. Par ailleurs, la norme Δ du flux optique est calculée pour chaque pixel ($\Delta = \sqrt{{\Delta_x}^2 + {\Delta_y}^2}$). Pour finir, l'image représentant la norme du flux optique a son histogramme égalisé de manière à mieux répartir les intensités sur l'ensemble de la plage de valeurs possibles. Le résultat est présenté sur la Figure 4.23b.



(a) Flux optique (b) Norme du flux optique

Figure 4.23 – Flux optique entre l'image téléobjectif et l'image grand-angle transformée dans l'espace géométrique de l'image téléobjectif.

Plus le niveau de gris est faible, plus le déplacement est faible, et donc plus la probabilité que le pixel appartienne à l'objet d'intérêt est élevée. Cette méthode permet d'obtenir une simili carte de profondeurs, que nous notons D_P.

Cependant, il arrive dans certains cas que l'appariement des points caractéristiques ne donne pas de bons résultats. En fait, dans le cas de petits fruits ou de petites fleurs, l'objet d'intérêt n'est qu'une petite proportion du cercle dans lequel sont conservés les points d'intérêt et une partie des points caractéristiques appartiennent donc en fait à l'arrière-plan. Il en résulte un recalage sur le fond et non sur l'objet d'intérêt (Figure 4.24).

Pour résoudre ce problème, un second recalage est réalisé. Notons H_1 la première matrice d'homographie calculée jusqu'à présent, et $telephoto_{H_1}$ le résultat de la transformation géométrique de l'image grand-angle par la matrice H_1. La seconde transformation géométrique va être réalisée entre l'image téléobjectif et $telephoto_{H_1}$. Comme précédem-

(a)

(b)

(c)

(d)

(e)

FIGURE 4.24 – Image grand-angle b) et téléobjectif c) d'un petit fruit, suscitant une détection et un appariement des points d'intérêt en partie sur le fond a). La matrice d'homographie étant incorrecte, la transformation géométrique de l'image grand-angle ne centre pas l'objet d'intérêt d), donnant donc une carte de profondeurs e) non pertinente.

ment, les points caractéristiques U-SURF sont détectés et calculés, mais cette fois avec un plus petit seuil de matrice Hessienne, noté σ_{H_2} (σ_{H_1} étant le seuil pour le premier recalage). Ceci permet d'extraire des points caractéristiques à l'intérieur de l'objet d'intérêt, même sur la surface lisse de certains fruits et pétales de fleurs, et pas seulement sur les contours. Deuxièmement, seuls les points caractéristiques dont les coordonnées se trouvent à l'intérieur de la trace utilisateur sont conservés. Au préalable, cette trace utilisateur a été dilatée de manière à inclure l'objet d'intérêt sans trop déborder sur le

fond. L'influence de la trace utilisateur est discutée dans la partie suivante. Étant donné que le nombre de points caractéristiques est élevé du fait du seuil σ_{H_2} faible, seules les 20 meilleures paires de points appareillés (ayant des similitudes les plus faibles) sont utilisées pour calculer la seconde matrice d'homographie H_2. Nous ne pouvons donc nous passer du premier recalage et considérer directement les points d'intérêt à l'intérieur de la trace car nous ne pouvons transformer la trace utilisateur dans l'espace géométrique de l'image grand-angle sans connaître la transformation géométrique entre les deux images. La Figure 4.25 montre les points caractéristiques extraits sur l'image téléobjectif originale et l'image $telephoto_{H_1}$, ainsi que les 20 paires de points conservées.

(a) $telephoto_{H_1}$ (b) Image téléobjectif (c) Appariement

Figure 4.25 – Appariement des points caractéristiques restreints sur la trace utilisateur.

À nouveau, la matrice H_2 est utilisée pour calculer la transformation géométrique sur $telephoto_{H_1}$ pour obtenir $telephoto_{H_2}$ ($telephoto_{H_2} = H_2 \cdot telephoto_{H_1}$). Les résultats sont présentés sur le Figure 4.26.

(a) Image téléobjectif (b) $telephoto_{H_1}$ (c) $telephoto_{H_2}$ (d) Flux optique

Figure 4.26 – Le calcul de la transformation géométrique entre les images a) et b) permet de générer l'image c) où l'objet d'intérêt est centré, permettant de calculer l'image de profondeur d) grâce à la norme du flux optique calculé entre a) et c).

Pour résumer, la Figure 4.27 présente un schéma du processus pour obtenir la carte de profondeurs D_P.

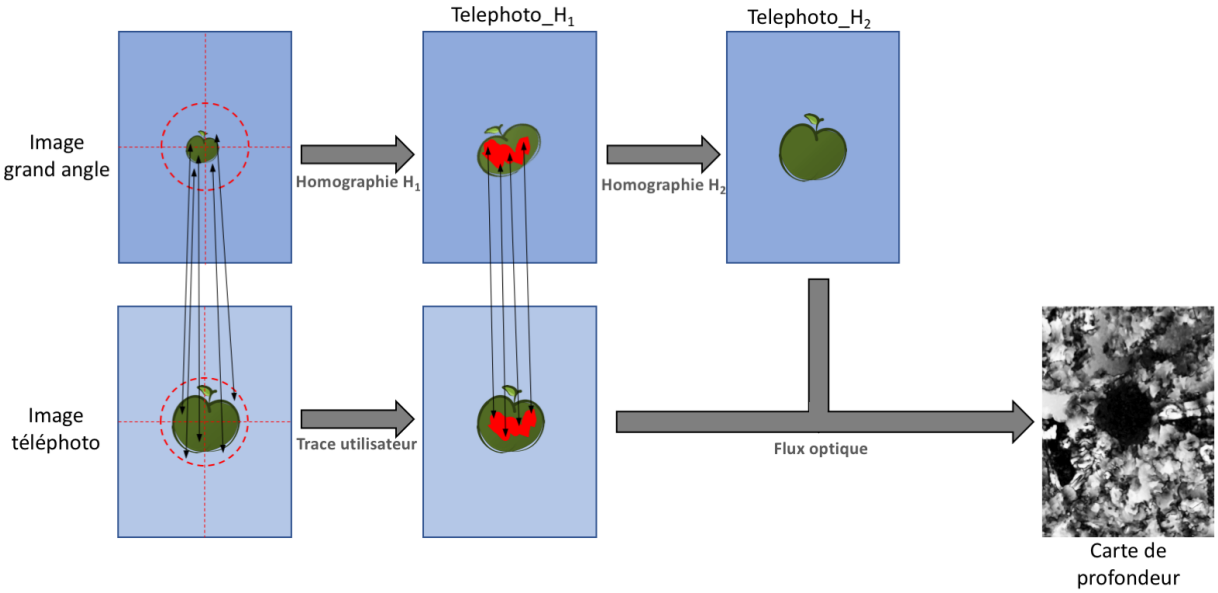


FIGURE 4.27 – Schéma de la génération de la carte de profondeurs D_P.

4.3.1 Influence de la trace utilisateur

Ici aussi, pour la construction de la carte de profondeurs, la trace utilisateur joue un rôle important lors du second recalage, pour s'assurer que celui-ci se fait bien uniquement sur l'objet d'intérêt dans le cas de petits fruits ou petites fleurs et non sur le fond (Figure 4.24). L'objectif de cette partie est donc d'étudier l'influence de la taille de la trace utilisateur sur la carte de profondeurs. Nous proposons pour cela de calculer la moyenne et l'écart-type de l'intensité des pixels de la carte de profondeurs à l'intérieur de l'objet d'intérêt grâce à la vérité terrain. Plus la carte de profondeurs est correcte, plus la zone de l'objet d'intérêt a un niveau de gris faible avec un faible écart-type.

De même que pour l'étude de l'influence de la trace utilisateur sur la carte de MLP, la trace est érodée ou dilatée avec un élément structurant circulaire ayant un diamètre de trois pixels. La trace utilisateur est érodée jusqu'à ce qu'elle soit réduite de $x\%$ par rapport à la trace initiale, ou est dilatée jusqu'à ce que celle-ci soit étendue de $x\%$ par rapport à la trace initiale. Les évolutions de la moyenne et de l'écart-type en fonction de la taille de la trace utilisateur sont présentées sur les Figures 4.28 et 4.29 respectivement.

La partie gauche de l'axe des abscisses correspond à l'érosion de la trace utilisateur, et la partie droite à sa dilatation. Nous pouvons voir qu'éroder la trace entraîne une légère augmentation du niveau de gris médian ainsi que de l'écart-type dans la zone de l'objet d'intérêt, indiquant une détérioration des cartes de profondeur. En effet, plus la trace est réduite, plus les points caractéristiques U-SURF utilisés pour calculer le second recalage sont dans une zone réduite de l'objet d'intérêt. Ainsi, seul le centre de la fleur ou du fruit est recalé, et non l'ensemble. À l'opposé, la dilatation fait baisser le niveau de gris médian au sein de l'objet d'intérêt ainsi que légèrement l'écart-type, jusqu'à une dilatation de 50%. Cependant, nous pouvons déduire de la moyenne des niveaux de gris maximum que pour

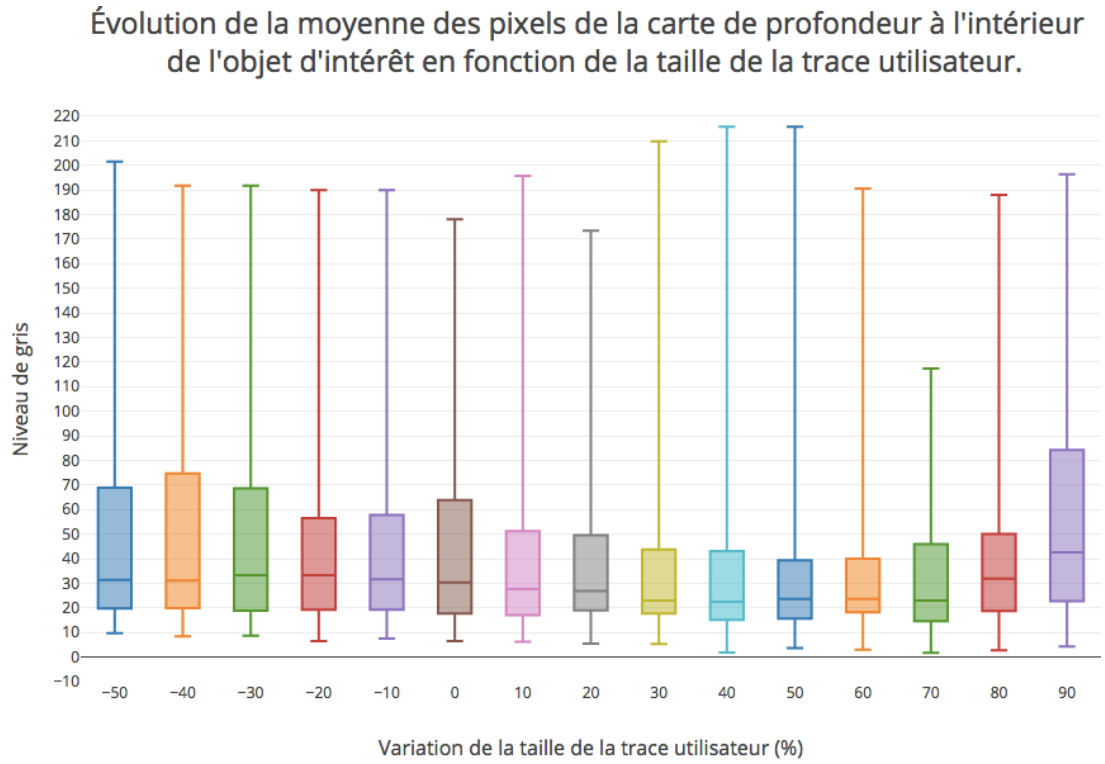


Figure 4.28 – Influence de la surface de la trace utilisateur sur la carte de profondeurs sur la moyenne des pixels de la carte de profondeurs à l'intérieur de la trace utilisateur.

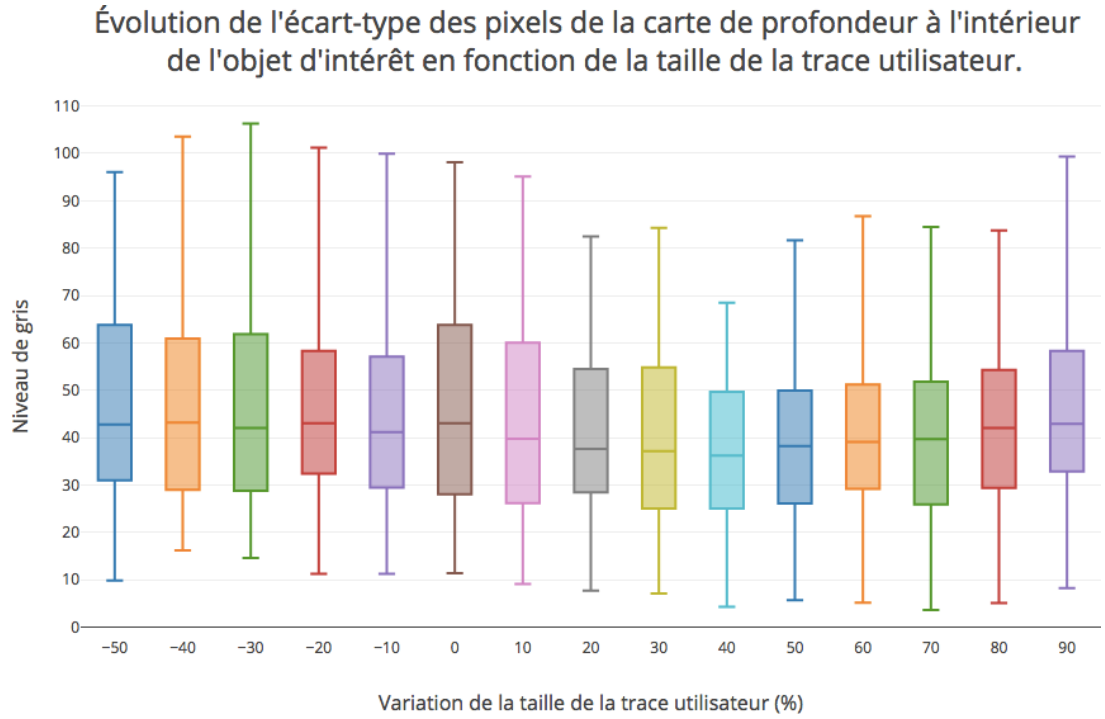


Figure 4.29 – Influence de la surface de la trace utilisateur sur la carte de profondeurs sur l'écart-type des pixels de la carte de profondeurs à l'intérieur de la trace utilisateur.

certaines images la dilatation détériore la qualité de la carte de profondeurs. En effet, la marque dessinée par l'utilisateur se situe au centre de l'objet à segmenter, et ne correspond

pas à son ensemble. Dilater cette marque permet d'extraire les points caractéristiques pas seulement dans le centre de l'objet, mais aussi dans la périphérie. Cette dilatation doit être effectuée dans la limite du raisonnable, car au-delà d'une certaine taille, la trace dilatée englobe aussi le fond de l'image. Cependant, la dilatation de la trace entraîne une augmentation des temps de calcul. En effet, plus la trace utilisateur est large, plus le nombre de points caractéristiques extraits est élevé, d'autant que le seuil σ_{H_2} est faible, et plus il y a d'appariements à effectuer. De ce fait, nous fixons la dilatation de la trace à 40% pour le calcul de l'image $telephoto_{H_2}$ de la carte de profondeurs. Avec une dilatation de la trace à 40%, la variation de l'écart-type est la plus faible, indiquant que les niveaux de gris de la carte de profondeurs sont homogènes à l'intérieur de la trace utilisateur, donc au niveau de l'objet d'intérêt. La Figure 4.28 nous indique que, dans la majorité des images, ce niveau de dilatation permet d'obtenir une carte de profondeurs avec une profondeur faible au niveau de l'objet d'intérêt, mais que pour certaines images la carte de profondeurs sera fausse (avec une profondeur élevée au niveau de l'objet d'intérêt). La prise en compte de la qualité des cartes de profondeur va être étudiée dans la partie 4.4.3.

Sur l'ensemble de notre base de données de 57 paires d'images la première homographie est calculée en moyenne en 0,69 seconde, la seconde homographie en 0,64 seconde et la carte des flux optiques en 0,09 seconde, avec un MacBook Air possédant un processeur Intel Core i7 de 2,2 GHz et une mémoire DDR3 de 8 Go 1600 MHz.

4.3.2 Construction de la carte avec un même appareil photo

Jusqu'à maintenant, les expérimentations ont été réalisées avec les photos prises avec un seul système d'acquisition (iPhone). La transformation géométrique entre les deux images est à peu près la même si l'utilisateur ne bouge pas trop entre les deux prises de vues. D'autre part, notre méthode nécessite théoriquement d'avoir un smartphone avec deux caméras en face arrière. Dans cette partie, nous allons voir s'il est possible de produire une carte de profondeurs avec deux photos prises avec une seule caméra avec un angle ou un zoom différent. Nous effectuons les tests avec un autre smartphone (Samsung Galaxy A3 SM-A300FU) et un appareil photo de type Reflex.

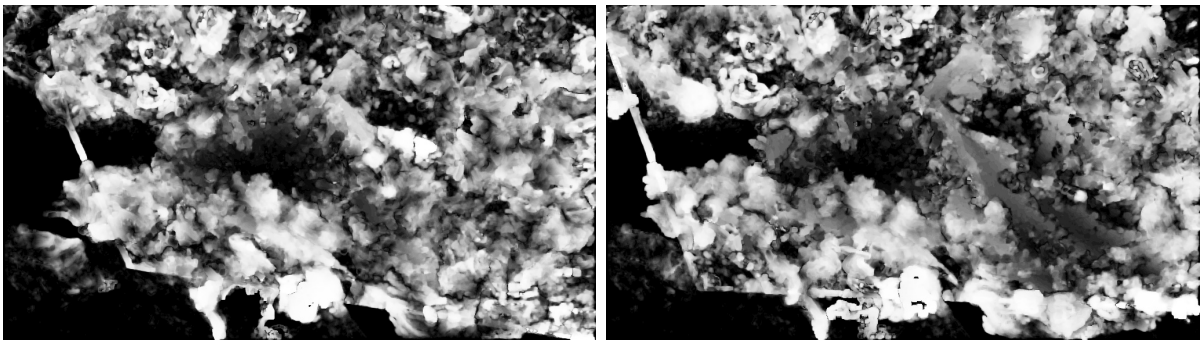
Changement d'angle de vue

Avec le second smartphone, nous testons si l'algorithme permettant de générer la carte de profondeurs peut donner de bons résultats en prenant en entrée deux photos de l'objet d'intérêt photographié avec la même caméra mais avec un angle de vue différent. Nous pouvons prendre la paire d'images montrée sur la Figure 4.30.

Nous ne pouvons pas directement calculer le flux optique entre ces deux images car il n'est pas garanti que l'objet soit au centre des deux images. Cependant, comme les deux images sont dans le même espace géométrique (la distance entre le smartphone et l'objet

Figure 4.30 – Paire d'images du gland du chêne chevelu (*Quercus cerris*) photographié sous deux angles légèrement différents.

d'intérêt est la même, seul l'angle change), un seul recalage peut être suffisant. Dans ce cas, nous pouvons directement considérer les points d'intérêt SURF qui sont uniquement dans la trace utilisateur, et non dans le centre de l'image. Si ces deux images sont fournies en entrée de l'algorithme d'élaboration de la carte de profondeurs, nous obtenons les résultats présentés sur la Figure 4.31.



(a) Carte de profondeurs avec deux recalages. (b) Carte de profondeurs avec un seul recalage.

Figure 4.31 – Cartes de profondeur obtenues pour la paire d'images de la Figure 4.30.

Dans un premier temps, nous pouvons remarquer que notre algorithme produit des cartes de profondeur cohérentes (avec un niveau de gris faible là où se situe l'objet d'intérêt) pour deux photos prises avec une seule caméra sur un téléphone portable sous deux angles de vue différents. Ensuite, nous pouvons remarquer que, dans cette configuration, il est possible de s'affranchir du premier recalage pour n'effectuer qu'un recalage uniquement à partir des points d'intérêts extraits dans la trace utilisateur.

Changement d'angle de champ

Plutôt que de reprendre en photo le fruit ou l'inflorescence sous un angle de vue différent, nous regardons s'il est possible d'effectuer un zoom avec le même appareil photo. Pour cela, nous utilisons un appareil photo de type reflex. L'objet d'intérêt est photographié à la même distance mais avec des focales différentes. Pour ces expériences, nous utilisons les photos de la Figure 4.32.

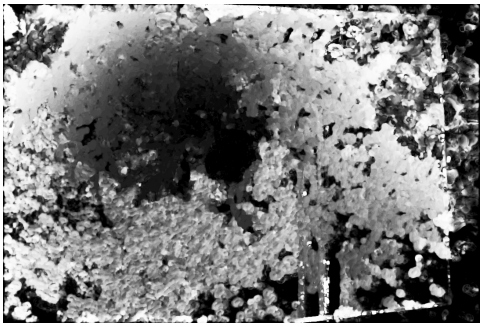
(a) (b)

(c) (d)

FIGURE 4.32 – Fleur de grenadier (*Punica granatum*) photographiée avec différentes focales.

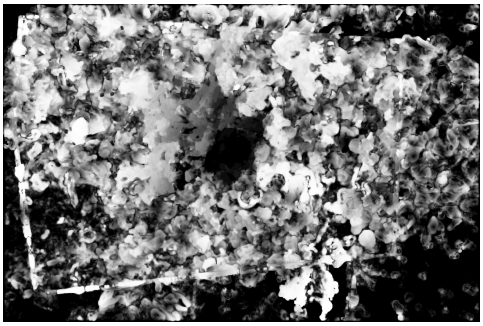
Nous choisissons de calculer les cartes de profondeur par rapport à l'image 4.32c. Nous utilisons le processus complet de génération de la carte de profondeurs comme le détaille la Figure 4.27. Premièrement, la carte de profondeurs est calculée entre les images 4.32c (image grand-angle) et 4.32d (image téléobjectif). Deuxièmement, nous inversons l'image grand-angle avec l'image téléobjectif pour vérifier que la méthode n'est pas mise en défaut. Ainsi la carte de profondeurs est calculée entre les images 4.32b et 4.32c, et entre les images 4.32a et 4.32c, ces deux dernières présentant une plus grande différence de focale entre elles. Pour finir, la carte de profondeurs est calculée entre les deux images les plus opposées en termes de focale, c'est-à-dire les images 4.32a et 4.32d. Les résultats sont visibles sur la Figure 4.33. L'image citée en premier est fournie en entrée de l'algorithme en tant qu'image grand-angle, et la seconde en tant qu'image téléobjectif.

La carte de profondeurs calculée entre les images 4.32b et 4.32c est proche de la vérité terrain. En effet, la fleur, qui est au premier plan correspond aux niveaux de gris les plus faibles au centre de la carte. Puis ce sont les feuilles juste derrière la fleur, correspondant au second plan, qui ont des niveaux de gris un peu plus élevés. Enfin, le reste de l'image correspondant à l'arrière-plan a des niveaux de gris élevés. Le changement d'angle de champ important entre les deux images permet de reconstruire de manière assez précise cette carte. Cependant, les cartes de profondeur entre les images 4.32b et 4.32c, et entre les images 4.32a et 4.32d, donnent des résultats plus exploitables pour une segmentation de l'objet, car uniquement les pixels à la place de l'objet d'intérêt ont un niveau de gris

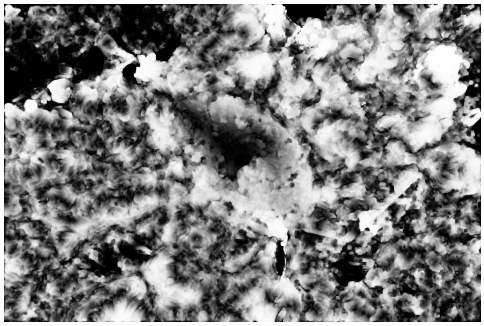


(a) Carte de profondeurs entre les images 4.32d et 4.32c.

(b) Carte de profondeurs entre les images 4.32b et 4.32c.



(c) Carte de profondeurs entre les images 4.32a et 4.32c.



(d) Carte de profondeurs entre les images 4.32a et 4.32d.

Figure 4.33 – Résultats de cartes de profondeur pour les images de la Figure 4.32.

faible et non le feuillage du second plan. Pour finir, la carte de profondeurs obtenue avec les deux images ayant la plus grande différence de focale donne aussi un bon résultat, avec un flux optique faible au niveau de la fleur. Il y a juste une partie du feuillage du second plan qui a des niveaux de gris faibles, comme l'objet d'intérêt, sur la carte de profondeurs. Enfin, la carte de profondeurs a été calculée entre les images 4.32a et 4.32b. Le résultat est présenté sur la Figure 4.34.

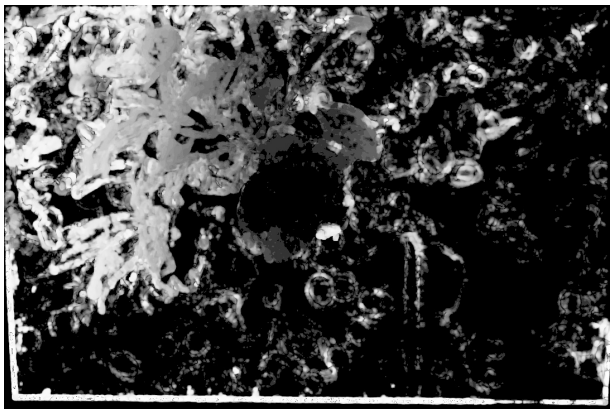


Figure 4.34 – Carte de profondeurs entre les images 4.32a et 4.32b.

Bien que les pixels de la carte de profondeurs aient des niveaux de gris faibles, et le feuillage au second plan des niveaux de gris plus élevés, l'arrière-plan présente des niveaux

de gris faible malgré la grande profondeur. Ceci ne vient pas des recalages mais du calcul du flux optique dans cette région. En effet, le flou important dans cette région induit une uniformisation des valeurs des pixels qui ne permet pas à l'algorithme de Farneback de faire une mise en correspondance correcte entre les pixels entre les deux images.

Par conséquent, il est possible de recréer des cartes de profondeur avec une seule caméra en faisant varier la focale. Il n'existe actuellement que très peu d'appareils photo sur smartphone sur lequel la focale est variable (Asus Zenfone Zoom ou Samsung Galaxy S4 Zoom). Cependant, cette caractéristique ne semble pour l'instant pas à la mode, contrairement à la double caméra en face arrière, à cause de la place importante que prennent les différents systèmes de lentilles. De plus, avec une grande focale l'arrière-plan risque d'être flou, impliquant de mauvais résultats lors du calcul du flux optique. Dans le cas où l'appareil de prise de vue permet de rendre l'arrière-plan flou, un autre algorithme pour la segmentation du premier plan serait plus approprié, tel que le travail de Saitoh *et al.* [SAK04].

4.4 Contour actif

Nous avons vu dans la partie 4.2 que la carte de MLP, notée D_{MLP}, donne de bons résultats, mais que cette méthode tend à sur-segmenter l'objet d'intérêt. Nous proposons donc d'utiliser la carte de profondeurs, notée D_P, afin de réduire cette sur-segmentation en restreignant la segmentation à l'objet d'intérêt au premier plan. La méthode choisie pour segmenter de manière optimale l'objet d'intérêt en combinant la carte de MLP et la carte de profondeurs est l'algorithme des contours actifs. Un modèle de contour actif, souvent appelé snakes, est une structure dynamique introduite par Kass *et al.* [KWT88] pour délimiter le contour d'un objet à partir d'une image 2D. Le contour actif est une courbe Γ déformable et minimisant l'énergie, influencée par des contraintes et des forces externes qui l'attirent vers les contours de l'objet et des forces internes qui résistent à sa déformation. Le but étant de minimiser l'énergie totale présente le long de la courbe. L'énergie est constituée de trois termes :

$$E(\Gamma) = E_{interne}(\Gamma) + E_{attache}(\Gamma) + E_{externe}(\Gamma) \tag{4.10}$$

Le premier terme est l'énergie interne, qui régit la forme du contour afin d'obtenir un contour plus ou moins lisse et élastique. En effet, le contour doit conserver une forme arrondie en minimisant les dérivées d'ordre 1, 2, *etc.* et doit empêcher un point du contour de se détacher trop loin du reste de la courbe. Idéalement, l'énergie interne est minimale pour un cercle où tous les points sont régulièrement espacés. Le deuxième terme est l'énergie d'attache aux données de l'image prenant en compte par exemple l'intensité des pixels ou dans la plupart des cas l'information de gradient pour que la courbe tende vers

les contours de l'objet à segmenter. Enfin, le troisième terme est l'énergie externe. Cette partie permet au concepteur de rajouter des contraintes en plus des données de l'image à segmenter.

Le modèle de contour actif mis en œuvre dans notre méthode est dérivé des travaux de Mille [Mil09], actuellement utilisé dans l'application *Folia* pour la segmentation des feuilles. Le contour actif est initialisé sur les bords de la trace utilisateur et se calcule sur l'image téléobjectif. Dans notre cas, l'énergie interne est constituée de deux termes :

$$E_{interne}(\Gamma) = E_{lissage}(\Gamma) + E_{ballon}(\Gamma) \tag{4.11}$$

L'énergie de lissage permet de caractériser le lissage de la courbe, et l'énergie de ballon traduit la force avec laquelle la courbe s'écarte de sa position initiale (la trace utilisateur) pour venir toucher les bords de l'objet d'intérêt. L'énergie d'attache aux données correspond aux gradients de l'image. Elle est faite pour pénaliser les points du contour qui se trouveraient sur des pixels de faible intensité de gradient. Enfin, l'énergie externe est également composée de deux termes :

$$E_{externe}(\Gamma) = E_{MLP}(\Gamma) + E_{profondeur}(\Gamma) \tag{4.12}$$

Ces deux énergies sont tirées des deux cartes précédemment détaillées. Le premier terme correspond à l'énergie dans le contour fermé de la carte de MLP (D_{MLP}) et le second à l'énergie dans le contour fermé de la carte de profondeurs (D_P) selon les équations suivantes :

$$E_{MLP}(\Gamma) = \int_{\Omega(\Gamma)} (C_{MLP}(x, y, \Omega_{trace}))dxdy \tag{4.13}$$

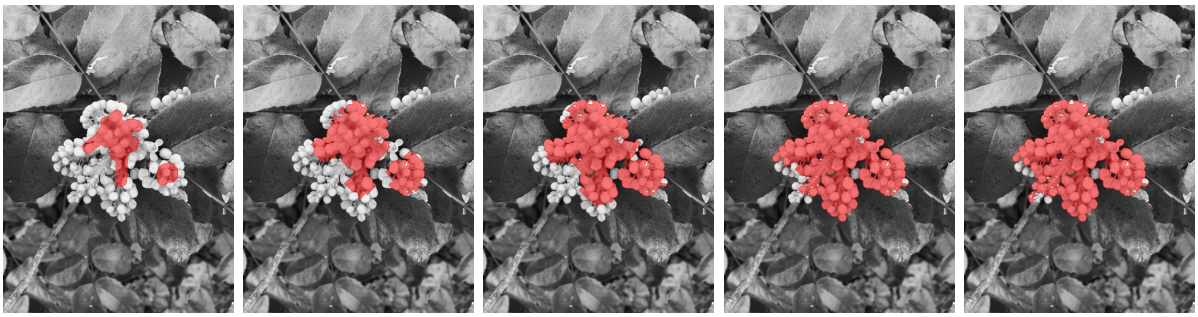
$$E_{profondeur}(\Gamma) = \int_{\Omega(\Gamma)} (C_P(x, y, \Omega_{trace}))dxdy \tag{4.14}$$

La carte de couleurs n'est pas prise en compte car l'information de couleur est intégrée dans la carte de MLP. D'autre part, il se peut que les différentes énergies n'aient pas la même importance pour le calcul du contour actif en ce qui concerne la segmentation de l'objet d'intérêt. Les diverses énergies qui interviennent dans l'évolution du contour forment un ensemble difficile à harmoniser en un tout cohérent, qui soit efficace quelle que soit l'image. Ainsi, le calcul de l'énergie est réalisé *via* l'équation suivante :

$$E(\Gamma) = \alpha E_{ballon}(\Gamma) + \beta E_{lissage}(\Gamma) + \gamma E_{gradient}(\Gamma) + \delta E_{MLP}(\Gamma) + \epsilon E_{profondeur}(\Gamma) \tag{4.15}$$

Dans la pratique, les différents coefficients α, β et γ ont été fixés empiriquement sans trop de variations par rapport aux coefficients de ces mêmes énergies ($E_{ballon}(\Gamma)$,

$E_{lissage}(\Gamma)$ et $E_{gradient}(\Gamma)$) utilisées pour la segmentation des feuilles dans l'application *Folia*. Les valeurs choisies sont $\alpha = 0.3$, $\beta = 6.0$ et $\gamma = 8.0$. Les poids δ et ϵ des deux cartes sont étudiés dans la partie 4.4.3. Le contour actif est calculé jusqu'à convergence de l'énergie $E(\Gamma)$. La Figure 4.35 montre l'évolution du contour actif sur un organe végétal.



(a) Initialisation (b) Itération 1 (c) Itération 2 (d) Itération 4 (e) Convergence

FIGURE 4.35 – Évolution du modèle de contour actif initialisé sur la trace utilisateur a). L'image est en niveau de gris pour une meilleure visualisation.

En premier lieu, les performances du modèle de contour actif sont comparées à la méthode CC_{MLP} sur la carte de MLP uniquement. En d'autres termes, le coefficient ϵ de l'équation 4.15 est mis à zéro afin que la carte de profondeurs ne soit pas prise en compte. Cette méthode est dénotée AC_{MLP}. Les mesures de qualité de segmentation (précision, rappel, Indice de Dice, Manhattan et Jaccad) sont présentées dans le Tableau 4.2.

	Précision	**Rappel**	**Dice**	**Manhattan**	**Jaccard**
CC_{MLP}	0.6519	0.9362	0.7192	0.9439	0.6163
AC_{MLP}	0.8569	0.8140	0.7960	0.9774	0.7000

TABLE 4.2 – Mesures de qualité de segmentation pour les méthodes CC_{MLP} et AC_{MLP} prenant en compte uniquement la carte de MLP calculée sur les images dites téléobjectifs.

Bien que la segmentation CC_{MLP} donne de bons résultats visuels comme l'illustre la Figure 4.15, AC_{MLP} améliore clairement la segmentation. Cela est dû au fait que CC_{MLP} tient également compte du gradient et des énergies internes (notamment le lissage). Le rappel et la précision montrent que l'objet d'intérêt est sur-segmenté avec les composantes connexes, par exemple lorsque des objets similaires se trouvent derrière l'objet d'intérêt. $E_{gradient}(\Gamma)$ empêche le contour de s'étendre à des objets similaires présents dans l'arrière-plan mais ce n'est pas suffisant.

4.4.1 Comparaison sur une base de données de feuille

Afin de comparer la segmentation réalisée avec la carte de MLP, cette carte a été calculée sur la base de données de feuilles utilisée par Grand-Brochier *et al.* dans leur étude comparative pour la segmentation des feuilles d'arbres [Gra+15]. La base de données est

composée d'images de feuilles photographiées en milieu naturel associées à une marque d'utilisateur. Selon cet article, la meilleure méthode est le contour actif guidé, notée GAC. C'est cette méthode qui est embarquée dans *Folia* pour réaliser la segmentation des feuilles. Le contour actif est "guidé" dans le sens où un modèle déformable de feuille est calculé en amont sur la feuille à segmenter, et que le contour ne peut trop s'écarter de ce modèle. Nos résultats sur cette base de données sont présentés dans le Tableau 4.3.

Table 4.3 – Comparaison des résultats de CC_{MLP} et (AC_{MLP} avec GAC pour la segmentation de feuilles.

	Précision	**Rappel**	**Dice**	**Manhattan**	**Jaccard**
GAC	**0.9879**	0.8281	**0.8934**	0.9309	**0.8191**
CC_{MLP}	0.6519	**0.9362**	0.7192	**0.9439**	0.6163
AC_{MLP}	0.8569	0.8140	0.7960	**0.9774**	0.7000

Le Tableau 4.3 montre que CC_{MLP} et AC_{MLP} surpassent légèrement d'un point de vue rappel la meilleure méthode de segmentation de feuilles de l'état de l'art (GAC) sans supposition *a priori* de forme concernant l'objet à segmenter. Comme le contour actif (AC_{MLP}) prend l'information du contour de la feuille (gradient), il a tendance à rester à l'intérieur de l'objet d'intérêt, comme le montrent les valeurs de précision. Ces deux méthodes sont cependant moins précises que l'algorithme GAC. Concernant la méthode GAC, l'information *a priori* de forme permet de contraindre correctement le contour actif dans la feuille.

4.4.2 Contribution de la carte de profondeurs

Dans le cas où deux fleurs ou fruits sont côte à côte, la carte de MLP n'est pas suffisante pour segmenter l'objet d'intérêt. La Figure 4.36 montre un exemple de cette configuration. Pour la segmentation des feuilles, cette configuration est traitée par l'utilisation du modèle déformable de feuille dans la méthode GAC. Pour rappel, la grande variété de formes de fleurs et de fruits ne permet pas l'utilisation d'un modèle pour contraindre la segmentation. C'est pourquoi la carte de profondeurs est utilisée afin de remplacer le modèle de forme. La segmentation réalisée avec le contour actif en prenant en compte la carte D_{MLP} et la carte D_P est notée $AC_{MLP,P}$.

La Figure 4.36e montre clairement la contribution de la carte de profondeurs. Grâce à la prise en compte de D_P le contour actif est contraint sur la fleur du premier plan, et ne déborde pas sur la seconde fleur comme la carte de MLP pourrait l'inciter.

4.4.3 Gestion des poids des deux cartes

Cependant, la difficulté est de déterminer les paramètres δ et ϵ des deux cartes utilisées pour l'évolution du contour actif. La Figure 4.37 montre l'évolution des mesures de qualité

(a) (b) (c) (d) (e)

FIGURE 4.36 – Contribution de la carte de profondeurs par rapport à la carte de MLP seule. a) Image initiale, b) Carte de profondeurs D_P, c) Carte de MLP D_{MLP}, d) Segmentation contour actif avec D_{MLP} seule, e) Segmentation contour actif avec D_{MLP} et D_P.

de segmentation en fonction de l'importance de la carte de profondeurs. Le paramètre ϵ varie entre 0 et 0.7. Le poids δ de la carte de MLP est quant à lui fixé à 0.4.

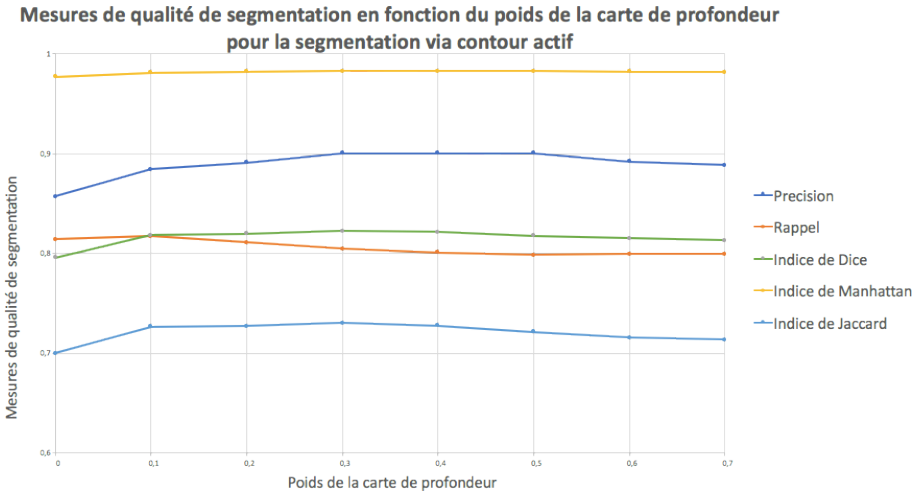


FIGURE 4.37 – Variation des indices de mesures de qualité de segmentation en fonction du poids ϵ de la carte de profondeurs D_P.

Ces résultats ne fournissent pas un rapport pertinent entre les poids des deux cartes. Parfois, D_{MLP} donne des informations plus pertinentes que D_P, ou vice versa. Pour résoudre ce problème, nous proposons de calculer les poids δ et ϵ de chaque carte en fonction de la moyenne des valeurs des pixels de la carte considérée à l'intérieur de la marque de l'utilisateur (équation 4.16), de telle sorte que si une erreur plus importante est commise par l'une des cartes, elle ait moins d'influence sur l'évolution du contour.

$$\delta = \frac{1}{2} - \frac{avg(D_{MLP} \in \Lambda)}{255} \quad \text{et} \quad \epsilon = \frac{1}{2} - \frac{avg(D_P \in \Lambda)}{255} \tag{4.16}$$

Les résultats de cette méthode sont présentés dans le Tableau 4.4. Une comparaison avec l'utilisation de la carte de MLP seulement (AC_{MLP}) est également présentée dans ce tableau.

Table 4.4 – Mesure de qualité de segmentation pour la segmentation via contour actif avec des poids δ et ϵ adaptés aux images.

	Précision	**Rappel**	**Dice**	**Manhattan**	**Jaccard**
AC_{MLP}	0.8569	0.8140	0.7960	0.9774	0.7000
$AC_{MLP,P}$ avec poids fixes	**0.9007**	0.8175	0.8226	0.9830	0.7304
$AC_{MLP,P}$ avec poids calculés	0.8967	**0.8292**	**0.8442**	**0.9838**	**0.7534**

Nous pouvons voir que cette méthode donne les meilleurs résultats par rapport au poids fixes, excepté pour la précision, mais avec une diminution de seulement 0,44%. Cependant, nous avons aussi vérifié que cette méthode permet d'avoir une meilleure précision de la segmentation qu'avec AC_{MLP}.

4.5 Conclusion

Pour l'identification des arbres et des plantes, les caractéristiques des fleurs et des fruits sont importantes pour les botanistes. Pour développer un algorithme de traitement d'images imitant le comportement du botaniste, il est nécessaire d'extraire l'organe du fond naturel complexe. La méthode proposée produit une segmentation efficace de l'objet d'intérêt à partir d'une double caméra de smartphone et avec une aide modeste de l'utilisateur (simple marque manuelle dans l'objet). Il a cependant été montré que la méthode développée pouvait fonctionner avec deux images prises à la même distance mais avec des angles de vue différents, ou avec deux images photographiées à la même distance mais avec des focales différentes.

La marque d'utilisateur est presque utilisée dans toutes les étapes du processus, jusqu'à la confiance des deux cartes dans l'algorithme de contour actif. L'influence de cette trace utilisateur a été discutée sur les différentes cartes. Comme notre méthode n'utilise pas d'informations *a priori* sur l'objet d'intérêt, nous pouvons aussi nous en servir pour segmenter d'autres objets comme nous l'avons montré pour les feuilles. La comparaison a été faite sur une base de données antérieurement utilisée dans la littérature et par l'application *Folia*.

Pour toutes les étapes de la segmentation proposée, les temps de calcul moyens sur notre base de données de fleurs et fruits photographiés avec l'iPhone 7 Plus sont les suivants : 0,69 seconde pour la première homographie et 0,64 seconde pour la seconde, 0,34 seconde pour le calcul de la carte des flux optiques, 0,09 seconde pour la génération de la carte des couleurs, 2,78 secondes pour la carte de MLP, et enfin 2,92 secondes pour le contour actif. Au total, la segmentation de l'objet est réalisée en 7,45 secondes en moyenne sur un MacBook Air possédant un processeur Intel Core i7 de 2,2 GHz et une mémoire DDR3 de 8 Go 1600 MHz.

Une fois l'objet d'intérêt segmenté, il s'agit d'en extraire des informations botaniques

de manière à pouvoir le reconnaître et ainsi identifier l'espèce végétale à laquelle il appartient.

Chapitre 5
Classification des fruits et fleurs

Ce chapitre expose les différents descripteurs utilisés pour classifier les fruits et les fleurs des arbres et arbustes présents en France métropolitaine. Les fruits et les inflorescences sont caractérisés après une étape de segmentation, comme celle proposée dans le chapitre précédent. Dans un premier temps, un rappel des contraintes botaniques est exposé, avant d'exposer des différents descripteurs, ainsi que le classifieur utilisé. La seconde partie sera consacrée à la comparaison des résultats de classification avec d'autres méthodes.

5.1 Cadre de travail

Les fruits, et surtout les fleurs, sont des organes botaniques portant énormément d'informations pour permettre l'identification de l'espèce. Ces organes sont cependant difficiles à étudier par les apprentis botanistes. En effet, pour rappel des novices n'arrivent pas toujours à identifier un organe comme une inflorescence ou comme un fruit. C'est pourquoi dans le chapitre précédent la méthode de segmentation ne prend déjà pas d'information *a priori* sur la nature de l'organe. Pour ce qui est de la reconnaissance des fruits et des fleurs, nous ne pouvons ici non plus attendre de l'utilisateur qu'il indique à l'application d'aide à l'identification des végétaux s'il a photographié une fleur ou un fruit. De ce fait, les descripteurs choisis devront permettre de retirer de l'information sur l'organe indépendamment de sa nature. Ainsi, nous ne pouvons reprendre des méthodes utilisées dans l'état de l'art pour la reconnaissance des fleurs telles que l'identification du nombre de pétales par exemple. De plus, le système de classification devra être robuste au nombre important de classes et à la variabilité de celles-ci.

5.1.1 Base de données

Dans un cadre idéal, la base de données d'images de fleurs et de fruits devrait être acquise avec un smartphone avec deux caméras en face arrière pour utiliser des images segmentées avec la méthode détaillée dans le chapitre précédent. Utiliser des organes segmentés permet d'extraire des descripteurs uniquement sur l'objet d'intérêt et non sur le fond et les autres éléments de l'image. Cependant, il n'a pas été possible de réaliser une base de données complète de fruits et de fleurs d'espèces d'arbres de France métropolitaine pour des raisons principalement de temps et de répartition des espèces en France. Le smartphone avec double caméra n'a été acquis qu'en milieu de thèse et il n'a pas été possible de photographier les fleurs et les fruits d'un même arbre puisque leurs apparitions sur la plante s'échelonnent tout au long de l'année avec des périodes de présence courtes (notamment pour les fleurs). Le fait de ne photographier que quelques individus ne peut représenter toute la variabilité des organes au sein de l'espèce. Par ailleurs, certaines espèces ne sont présentes que dans certaines régions de France. Enfin, avoir un nombre

réduit de photographes ajoute un biais dans la base de données. Pour ces raisons, ce sont de nouveau des images issues de la base de données ImagesCLEF qui sont utilisées dans ce chapitre. Ainsi les fruits et les fleurs sont photographiés par un grand nombre de personnes différentes dans des conditions de prises de vues variables. De plus, le nombre d'espèces ayant des photos de fleurs et de fruits est élevé.

L'application *Folia* reconnaît 120 espèces d'arbres. Dans le but d'augmenter les taux de d'identification de ces espèces à partir d'information des autres organes, la base de données que nous allons utiliser est un sous-ensemble de ces espèces issues de la base de données d'imageCLEF. Dans le cadre d'une identification utilisant conjointement les feuilles et les fruits ou les fleurs, nous ne pouvons considérer l'ensemble des 120 espèces car pour certaines de ces espèces des données sont manquantes, c'est-à-dire que pour certaines espèces, nous n'avons que des images de feuilles, mais pas de fleurs, ou inversement. De ce fait, le nombre d'espèces que nous pouvons considérer est réduit à 72 espèces. D'autre part, il se peut que, de manière exceptionnelle, il n'y ait pas d'images de fleur ou de fruit pour une espèce dans notre sélection d'espèces. C'est le cas notamment pour le figuier (*Ficus carica*) où le fruit (au sens gastronomique du terme) est en fait une fleur repliée sur elle-même dans laquelle les fruits se développent (voir Figure 5.1).

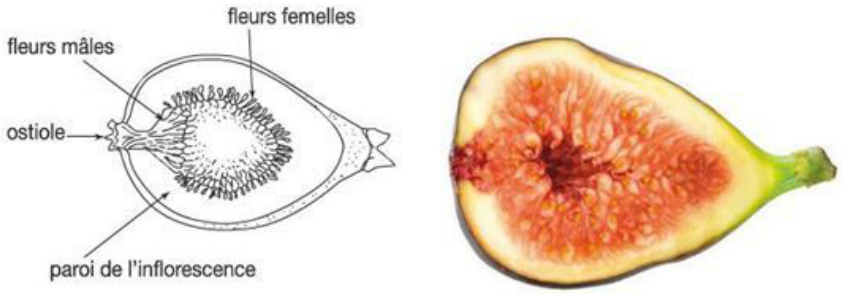


Figure 5.1 – Schéma d'une figue.

5.1.2 Préparation des données

Avant d'extraire des descripteurs, les images d'inflorescences et de fruits ont été segmentées. Les organes n'ont pas été segmentés à la main de manière à ce que les données ne soient pas trop "parfaites", le but étant que le système de classification soit robuste à des images segmentées *via* la méthode de segmentation présentée dans le chapitre précédent. Les images extraites de la base de données d'imageCLEF ont donc été segmentées de la manière suivante :

— Une trace utilisateur est faite sur l'objet d'intérêt et la carte de MLP est calculée. La composante connexe intersectant la trace utilisateur est conservée (méthode MLP_{CC}).

— Du fait qu'il n'y ait pas de carte de profondeur pour restreindre la sur-segmentation en cas d'objet similaire dans l'arrière-plan au contact de l'objet d'intérêt, la seg-

mentation est retouchée "à la main" de manière à ne garder que la fleur ou le fruit d'intérêt.

En tout, la base de données utilisée est composée de 3617 images de fruits ou de fleurs segmentées. Pour chaque classe de cette base, 2/3 des données sont prises pour l'apprentissage (2453 images), et 1/3 pour les tests (1164 images).

5.1.3 Fusion des inflorescences et des fruits

Comme indiqué précédemment, les fruits et les inflorescences ne sont pas traités séparément. Cependant les fleurs et les fruits appartenant à une espèce ne vont pas être considérés comme appartenant à la même classe : la variabilité intra-classe risquerait d'être trop grande pour la plupart des espèces qui ont une fleur et un fruit bien distincts. Ainsi, nous considérons dans la plupart des cas deux classes par espèce : "$fruitEspeceX$" et "$fleurEspeceX$". En tout, nous avons donc 136 classes de fruits ou fleurs pour 72 espèces. Pour rappel, le genre est le premier rang taxonomique après l'espèce qui traduit une proximité entre les espèces végétales. Au sein d'un même genre, les espèces produisent les mêmes fruits. Le rang taxonomique supérieur est la famille. Au sein d'une même famille les fleurs possèdent les mêmes propriétés morphologiques. Cette hiérarchie entre les espèces pourra être prise en compte pour l'identification des espèces d'arbre à partir des fruits et des fleurs. À titre d'information, la Figure 5.2 présente les différentes classes utilisées avec les correspondances pour les espèces, les genres et les familles.

5.2 Descripteurs

Pour décrire les fruits et les fleurs, nous ne pouvons nous inspirer directement des clefs botaniques utilisées par les experts pour identifier une espèce végétale, car celles-ci nécessitent régulièrement une dissection de l'organe pour connaître par exemple les positions relatives de l'ovaire ou des sépales pour la fleur, ou le nombre d'amandes à l'intérieur du fruit. À la place, nous nous inspirons des outils d'identification pour novices se basant sur des critères plus simples et intuitifs que sont la couleur, la forme et la texture.

5.2.1 Descripteur de couleur

Une caractéristique importante pour la caractérisation des fleurs dans les clefs botaniques pour novices est la couleur, malgré le fait que cela ne soit pas un critère botanique car la couleur des fleurs peut changer en fonction de la variété (exemple du laurier-rose *Nerium oleander*). En effet, comme dans l'application pour smartphone *Quelle est cette fleur ?* (voir Chapitre 1), les fleurs sont classées selon leur teinte (blanche, jaune, rouge, bleue, verte ou marron). Pour les fruits, la couleur peut aussi apporter des informations

Famille	Genre	Espece	Classe
Adoxaceae	Viburnum	Viburnum lantana L,	Flowers Viburnum lantana L,
			Fruits Viburnum lantana L,
		Viburnum opulus L,	Flowers Viburnum opulus L,
			Fruits Viburnum opulus L,
		Viburnum tinus L,	Flowers Viburnum tinus L,
			Fruits Viburnum tinus L,
Altingiaceae	Liquidambar	Liquidambar styraciflua L,	Flower Liquidambar styraciflua L,
			Fruit Liquidambar styraciflua L,
Anacardiaceae	Cotinus	Cotinus coggygria Scop,	Flowers Cotinus coggygria Scop,
			Fruits Cotinus coggygria Scop,
Apocynaceae	Nerium	Nerium oleander L,	Flowers Nerium oleander L,
			Fruits Nerium oleander L,
Aquifoliaceae	Ilex	Ilex aquifolium L,	Flowers Ilex aquifolium L,
			Fruits Ilex aquifolium L,
Araliaceae	Hedera	Hedera helix L,	Flowers Hedera helix L,
			Fruits Hedera helix L,
Betulaceae	Alnus	Alnus cordata (Loisel.) Duby	Fruits Alnus cordata (Loisel.) Duby
		Alnus glutinosa (L.) Gaertn,	Flowers Alnus glutinosa (L.) Gaertn,
			Fruits Alnus glutinosa (L.) Gaertn,
	Betula	Betula pendula Roth	Flowers Betula pendula Roth
			Fruits Betula pendula Roth
		Betula pubescens Ehrh,	Flowers Betula pubescens Ehrh,
	Carpinus	Carpinus betulus L,	Flowers Carpinus betulus L,
			Fruits Carpinus betulus L,
	Corylus	Corylus avellana L,	Flowers Corylus avellana L,
			Fruits Corylus avellana L,
	Ostrya	Ostrya carpinifolia Scop,	Flowers Ostrya carpinifolia Scop,
			Fruits Ostrya carpinifolia Scop,
Buxaceae	Buxus	Buxus sempervirens L,	Flowers Buxus sempervirens L,
			Fruits Buxus sempervirens L,
Cannabaceae	Celtis	Celtis australis L,	Flowers Celtis australis L,
			Fruits Celtis australis L,
Celastraceae	Euonymus	Euonymus europaeus L,	Flowers Euonymus europaeus L,
			Fruits Euonymus europaeus L,
Cornaceae	Cornus	Cornus mas L,	Flowers Cornus mas L,
			Fruits Cornus mas L,
		Cornus sanguinea L,	Flowers Cornus sanguinea L,
			Fruits Cornus sanguinea L,
Ebenaceae	Diospyros	Diospyros kaki L,f,	Flowers Diospyros kaki L,f,
			Fruits Diospyros kaki L,f,
Ericaceae	Arbutus	Arbutus unedo L,	Flowers Arbutus unedo L,
			Fruits Arbutus unedo L,
Euphorbiaceae	Euphorbia	Euphorbia characias L,	Flower Euphorbia characias L,
			Fruit Euphorbia characias L,
Fabaceae	Cercis	Cercis siliquastrum L,	Flowers Cercis siliquastrum L,
			Fruits Cercis siliquastrum L,
Fagaceae	Castanea	Castanea sativa Mill,	Flowers Castanea sativa Mill,
			Fruits Castanea sativa Mill,
	Fagus	Fagus sylvatica L,	Flowers Fagus sylvatica L,
			Fruits Fagus sylvatica L,
	Quercus	Quercus cerris L,	Fruits Quercus cerris L,
		Quercus coccifera L,	Flowers Quercus coccifera L,
			Fruits Quercus coccifera L,
		Quercus ilex L,	Flowers Quercus ilex L,
			Fruits Quercus ilex L,
		Quercus petraea Liebl,	Flowers Quercus petraea Liebl,
			Fruits Quercus petraea Liebl,
		Quercus pubescens Willd,	Flowers Quercus pubescens Willd,
			Fruits Quercus pubescens Willd,
		Quercus robur L,	Flowers Quercus robur L,
			Fruits Quercus robur L,
		Quercus rubra L,	Fruits Quercus rubra L,
Lauraceae	Laurus	Laurus nobilis L,	Flowers Laurus nobilis L,
			Fruits Laurus nobilis L,
Lythraceae	Punica	Punica granatum L,	Flowers Punica granatum L,
			Fruits Punica granatum L,
Magnoliaceae	Magnolia	Magnolia grandiflora L,	Flowers Magnolia grandiflora L,
			Fruits Magnolia grandiflora L,

Famille	Genre	Espece	Classe
Malvaceae	Tilia	Tilia cordata Mill,	Flowers Tilia cordata Mill,
			Fruits Tilia cordata Mill,
		Tilia platyphyllos Scop,	Flowers Tilia platyphyllos Scop,
			Fruits Tilia platyphyllos Scop,
Moraceae	Broussonetia	Broussonetia papyrifera (L,) Vent,	Flowers Broussonetia papyrifera (L,) Vent,
			Fruits Broussonetia papyrifera (L,) Vent,
	Ficus	Ficus carica L,	Fruits Ficus carica L,
	Morus	Morus alba L,	Flowers Morus alba L,
			Fruits Morus alba L,
Oleaceae	Ligustrum	Ligustrum vulgare L,	Flowers Ligustrum vulgare L,
			Fruits Ligustrum vulgare L,
	Olea	Olea europaea L,	Flowers Olea europaea L,
			Fruits Olea europaea L,
	Syringa	Syringa vulgaris L,	Flowers Syringa vulgaris L,
			Fruits Syringa vulgaris L,
Pittosporaceae	Pittosporum	Pittosporum tobira (Thunb,) W,T,Aiton	Flowers Pittosporum tobira (Thunb,) W,T,Aiton
			Fruits Pittosporum tobira (Thunb,) W,T,Aiton
Platanaceae	Platanus	Platanus x hispanica Mill, ex Mùnchh,	Flowers Platanus x hispanica Mill, ex Mùnchh,
			Fruits Platanus x hispanica Mill, ex Mùnchh,
Rhamnaceae	Frangula	Frangula dodonei Ard,	Flower Frangula dodonei Ard,
			Fruit Frangula dodonei Ard,
	Rhamnus	Rhamnus cathartica L,	Flowers Rhamnus cathartica L,
			Fruits Rhamnus cathartica L,
Rosaceae	Crataegus	Crataegus azarolus L,	Flowers Crataegus azarolus L,
			Fruits Crataegus azarolus L,
		Crataegus laevigata (Poir,) DC,	Flowers Crataegus laevigata (Poir,) DC,
			Fruits Crataegus laevigata (Poir,) DC,
		Crataegus monogyna Jacq,	Flowers Crataegus monogyna Jacq,
			Fruits Crataegus monogyna Jacq,
	Eriobotrya	Eriobotrya japonica (Thunb,) Lindl,	Flower Eriobotrya japonica (Thunb,) Lindl,
			Fruit Eriobotrya japonica (Thunb,) Lindl,
	Malus	Malus sylvestris Mill,	Flowers Malus sylvestris Mill,
			Fruits Malus sylvestris Mill,
	Prunus	Prunus avium (L,) L,	Flowers Prunus avium (L,) L,
			Fruits Prunus avium (L,) L,
		Prunus dulcis (Mill,) D,A,Webb	Flowers Prunus dulcis (Mill,) D,A,Webb
			Fruits Prunus dulcis (Mill,) D,A,Webb
		Prunus laurocerasus L,	Flowers Prunus laurocerasus L,
			Fruits Prunus laurocerasus L,
		Prunus padus L,	Flowers Prunus padus L,
			Fruits Prunus padus L,
		Prunus spinosa L,	Flowers Prunus spinosa L,
			Fruits Prunus spinosa L,
	Sorbus	Sorbus aria (L,) Crantz	Flowers Sorbus aria (L,) Crantz
			Fruits Sorbus aria (L,) Crantz
		Sorbus torminalis (L,) Crantz	Flowers Sorbus torminalis (L,) Crantz
			Fruits Sorbus torminalis (L,) Crantz
Salicaceae	Populus	Populus alba L,	Flowers Populus alba L,
			Fruits Populus alba L,
		Populus nigra L,	Flowers Populus nigra L,
			Fruits Populus nigra L,
		Populus tremula L,	Flowers Populus tremula L,
			Fruits Populus tremula L,
	Salix	Salix alba L,	Flowers Salix alba L,
			Fruits Salix alba L,
		Salix caprea L,	Flowers Salix caprea L,
			Fruits Salix caprea L,
Sapindaceae	Acer	Acer campestre L,	Flowers Acer campestre L,
			Fruits Acer campestre L,
		Acer monspessulanum L,	Flowers Acer monspessulanum L,
			Fruits Acer monspessulanum L,
		Acer opalus Mill,	Flowers Acer opalus Mill,
			Fruits Acer opalus Mill,
		Acer platanoides L,	Flowers Acer platanoides L,
			Fruits Acer platanoides L,
		Acer pseudoplatanus L,	Flowers Acer pseudoplatanus L,
			Fruits Acer pseudoplatanus L,
		Acer saccharinum L,	Fruit Acer saccharinum L,
Ulmaceae	Ulmus	Ulmus minor Mill,	Flowers Ulmus minor Mill,
			Fruits Ulmus minor Mill,

Figure 5.2 – Correspondances entre les différentes classes des organes et les espèces, genres et familles associés.

bien qu'elle soit moins pertinente à cause de la variation de teinte quand le fruit arrive à maturité.

De même que pour l'écorce, l'information de couleur doit être indépendante aux conditions de prise de vue. Aussi, le vecteur caractéristique doit être d'une taille raisonnable. Nous avons donc de nouveau choisi l'espace colorimétrique HSV et conservé uniquement l'histogramme H (la teinte) de taille 180 pour caractériser la couleur des fruits. L'espace L*a*b* n'a pas été testé car cette fois-ci toutes les couleurs sont potentiellement présentes sur les fleurs et les fruits et le canal H de l'espace HSV encode la couleur sur une taille

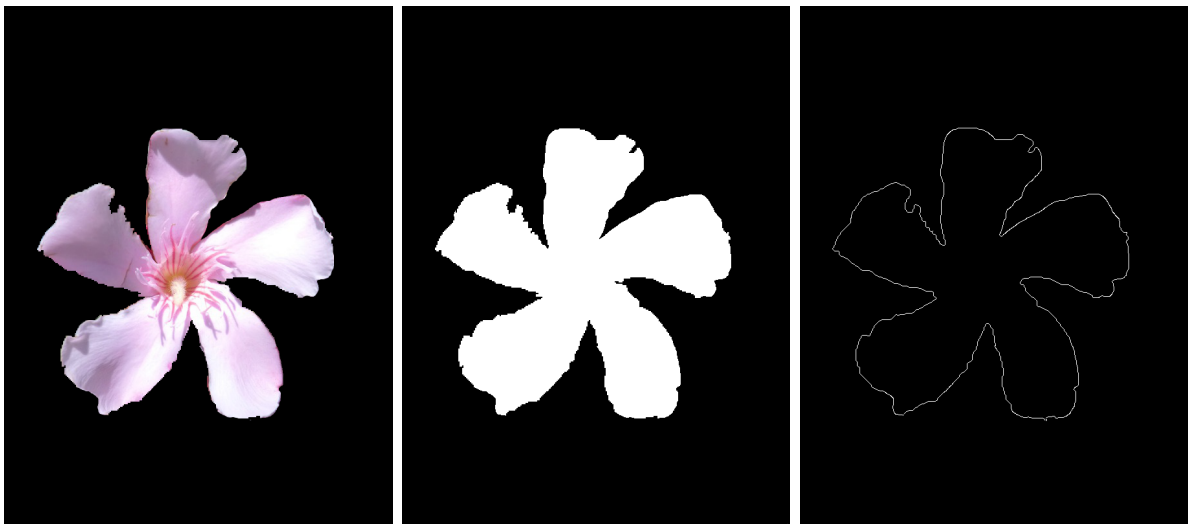
plus faible que les a* et b*.

Toutefois, la réduction de l'histogramme H est étudiée dans la partie 5.3.3.

5.2.2 Descripteur de forme

La forme de l'organe étant importante, nous avons considéré un descripteur de forme pour permettre de distinguer si le fruit est rond (drupe ou baie), allongé ou non (chaton, gousse), ou avec plusieurs parties (grappe de fleurs ou de fruits), ou pour représenter le nombre de pétale pour les fleurs en ayant. La forme peut être facilement caractérisée dans notre cas en récupérant son contour, l'objet d'intérêt étant supposé segmenté. Pour décrire la forme des objets binaires, il existe des paramètres génériques tels que le diamètre, l'aire, le périmètre, la concavité, *etc.* Mais ceux-ci ne sont pas toujours discriminants. De plus, sans contraintes de prise de vue, les descripteurs doivent être invariants à l'échelle, la rotation et la translation.

Les moments invariants de Hu [Hu62] sont des descripteurs de formes notables qui sont conformes avec ces contraintes. Ils sont notamment utilisés pour la reconnaissance de fruits maraîchers [Gar+15], mais n'ont jamais été testés sur des fruits d'arbres d'ornement. Les moments invariants de Hu sont extraits sur le contour d'une forme. La Figure 5.3 montre les différents traitements pour obtenir les contours. L'image en entrée est l'organe segmenté avec le fond noir. Un seuillage basique permet d'obtenir une image binaire de l'objet d'intérêt. Par la suite, le contour est extrait grâce à l'algorithme de suivi de contours de Suzuki *et al* [Suz+85]. Ce contour correspond à un vecteur de points.



(a) (b) (c)

FIGURE 5.3 – Les différents traitements avant l'extraction des moments de Hu avec l'image initiale a), l'image seuillée b) et les contours extraits c).

En traitement d'images, nous appelons moment d'une image une certaine moyenne pondérée (moment) des intensités des pixels de l'image, ou une fonction de ces moments,

habituellement choisie pour avoir une propriété ou une interprétation attrayante. Pour une image dont les intensités des pixels sont notées $I(x, y)$, le moment d'ordre $(p + q)$ de l'image I est calculé à l'aide de l'équation suivante :

$$M_{p,q} = \sum_x \sum_y x^p y^q I(x, y) \tag{5.1}$$

Cependant, le temps calcul des moments se révèle très chronophage à cause du nombre significatif d'additions et de multiplications. Une solution proposée par Li *et al.* est de calculer les moments en utilisant le théorème de Green [LS91]. D'après la démonstration de cet article, le moment d'ordre $(p + q)$ est calculé comme suit :

$$M_{pq} = \frac{1}{p+1} u_{pq} = -\frac{1}{q+1} v_{pq} \tag{5.2}$$

$$\text{avec} \quad u_{pq} = \sum_{(x_i, y_i \in \Gamma)} x_i^{p+1} y_i^q \quad \text{et} \quad v_{pq} = \sum_{(x_i, y_i \in \Gamma)} x_i^p y_i^{q+1} \quad \text{pour le contour } \Gamma \tag{5.3}$$

Les moments jusqu'à l'ordre $p + q \leq 3$ sont d'autre part définis comme suit :

$$\begin{aligned}
\mu_{00} &= M_{00} \\
\mu_{10} &= 0 \\
\mu_{01} &= 0 \\
\mu_{20} &= M_{20} - \overline{x} M_{10} \\
\mu_{02} &= M_{02} - \overline{y} M_{01} \\
\mu_{11} &= M_{11} - \overline{y} M_{10} \\
\mu_{30} &= M_{30} - 3\overline{x} M_{20} + 2\overline{x}^2 M_{10} \\
\mu_{12} &= M_{12} - 2\overline{y} M_{11} - \overline{x} M_{02} + 2\overline{y}^2 M_{10} \\
\mu_{21} &= M_{21} - 2\overline{x} M_{11} - \overline{y} M_{20} + 2\overline{y}^2 M_{01} \\
\mu_{03} &= M_{03} - 3\overline{y} M_{02} + 2\overline{y}^2 M_{01}
\end{aligned} \tag{5.4}$$

L'invariance à l'échelle peut être obtenue par normalisation. Les moments centraux sont alors définis comme suit :

$$\eta_{pq} = \frac{\mu_{pq}}{{\mu_{00}}^{\omega}} \quad avec \quad \omega = 1 + \frac{p+q}{2} \tag{5.5}$$

Comme le montrent les travaux de Hu *et al.* [Hu62], des moments invariants à la translation, à l'échelle et à la rotation peuvent être construits de la manière suivante :

$$\phi_1 = \eta_{20} + \eta_{02} \tag{5.6}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{5.7}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{5.8}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \tag{5.9}$$

$$\begin{aligned} \phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \end{aligned} \tag{5.10}$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} - \eta_{12})^2 - (\eta_{21} - \eta_{03})^2\right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{5.11}$$

$$\begin{aligned} \phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ -(\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \end{aligned} \tag{5.12}$$

L'ensemble des moments invariants de Hu constitue un vecteur caractéristique qui peut être utilisé pour caractériser la forme du fruit ou de l'inflorescence. Notons que celui-ci est de taille assez petite (7 moments).

5.2.3 Descripteur de texture

Pour finir, un élément régulièrement utilisé pour la classification des fruits ou des fleurs en traitement d'images est la texture [Zha+14 ; NZ06]. L'information de texture permet de différencier une fleur blanche unique avec des pétales lisses d'une inflorescence constituée d'une multitude de petites fleurs blanches par exemple (Figure 5.4). Pour cela, les méthodes basées sur les LBP (*Local Binary Patterns*) [OPM02], qui étiquettent les pixels d'une image en examinant le voisinage de chaque pixel et considèrent le résultat comme un nombre binaire, sont utilisées. L'avantage des LBP est leur robustesse aux changements de luminosité, ainsi que leur simplicité de calcul. Ce descripteur de texture est régulièrement utilisé dans la littérature, notamment pour les organes végétaux [Bak+13 ; DJ12 ; NLB12].

La première étape de la construction du descripteur de texture LBP est de convertir l'image en niveaux de gris. Pour chaque pixel de l'image en niveaux de gris, un voisinage de taille r (généralement égal à huit) entourant le pixel central est sélectionné. La valeur du pixel central est ensuite comparée à la valeur de chaque pixel du voisinage de r pixels (différence). Si la différence est positive ou égale à zéro, alors la valeur est fixée à 1,

(a) Fleur d'*Amelanchier ovalis*. (b) Fleur de *Sambucus nigra*.

Figure 5.4 – Deux inflorescences de même couleur pour lesquelles l'information de texture pourrait les différencier.

sinon à 0. L'ensemble de ces 1 et 0 forment un code binaire, qui après pondération donne une valeur numérique. Avec un voisinage de 3x3 (8 pixels), il y a un total de $2^8 = 256$ combinaisons possibles de codes binaires. À partir de là, la valeur LBP pour le pixel central est calculée. Les résultats de ce test binaire sont stockés dans un tableau de 8 bits, converti ensuite en décimal. La Figure 5.5 illustre les différentes étapes pour calculer la valeur du LBP pour un pixel.

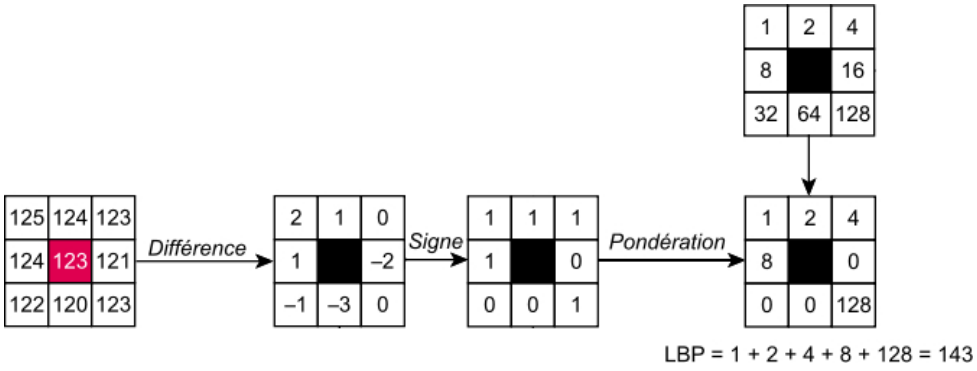


Figure 5.5 – Calcul du LBP pour le pixel central (en rouge).

Les valeurs de LBP sont calculées pour chaque pixel de l'image. Enfin, pour caractériser la texture d'une image, l'histogramme est calculé sur l'image de LBP obtenue. Avec un voisinage de taille 8, les valeurs de LBP s'échelonnent entre 0 et 255, permettant de construire un vecteur de taille 256.

Toutefois, l'utilisation de l'histogramme LBP donne un vecteur de taille assez grande, alors que certains motifs binaires ne sont que très peu représentés [OPM02]. Les motifs binaires avec plus de deux transitions de 0 à 1 ou 1 à 0 sont extrêmement rares dans les images de texture. Ces modèles sont appelés modèles "non uniformes", par opposition aux modèles comportant au plus deux transitions, appelés modèles "uniformes".

Par exemple, les motifs tels que '11000000' et '00001111' n'ont qu'une seule transition. De même, '11100111' et '00110000' ont deux transitions. Ces quatre modèles sont des exemples de modèles "uniformes". À l'opposé, '00110011' et '010101011' ont trois et sept transitions respectivement. Il s'agit de modèles "non uniformes". Pour un voisinage de taille 8 avec 256 motifs binaires de base, 58 seulement sont "uniformes". Les autres sont "non-uniformes", et apparaissent rarement dans les textures naturelles. Dans l'algorithme des LBP uniformes (U-LBP), seuls les 58 motifs "uniformes" ont un bin unique dans l'histogramme alors que tous les autres motifs "non-uniformes" sont rassemblés dans un seul bin. La taille de l'histogramme des LBP uniformes est de 59 bins. C'est cet histogramme qui est utilisé pour caractériser la texture des fruits et des fleurs.

5.3 Classifieur

Une fois les descripteurs de couleur (dimension 180), de forme (dimension 7) et de texture (dimension 59) calculés, ceux-ci sont concaténés pour former un seul vecteur caractéristique sur lequel va se baser le classifieur pour déterminer à quelle classe, et ensuite à quelle espèce la fleur ou le fruit appartient. Deux systèmes de classification ont été employés : un classifieur SVM et un classifieur MLP.

5.3.1 Augmentation des données

Pour être performants, les réseaux de neurones requièrent un nombre de données suffisant afin de calculer les poids de chaque neurone entre les différentes couches. Les espèces qui ont été sélectionnées ne présentent pas toutes un nombre suffisant d'images pour l'apprentissage. En effet, certaines espèces n'ont que deux images pour l'apprentissage des fruits ou des fleurs, tandis que d'autres en ont plus de 50. En moyenne, la population des classes est de 17 images (Figure 5.6). Avoir des classes équilibrées permet également au SVM de séparer plus efficacement les classes lors de la phase d'apprentissage.

Afin de pallier ces problèmes, nous proposons de faire une augmentation de données. L'augmentation d'images permet de créer artificiellement des images d'apprentissage à partir d'une base peu fournie grâce à des combinaisons de traitement d'images telles que des rotations, des décalages, des ajouts de bruits, *etc.* Nos images d'apprentissage sont augmentées grâce à de l'ajout de bruit gaussien sur les trois canaux RGB et grâce à une transformation de perspective. Nous n'effectuons pas de rotations ou de décalages car ces transformations n'influencent pas les caractéristiques des fruits et des fleurs que définissent les descripteurs choisis (forme, couleur, texture). En effet, les moments de Hu sont invariants aux rotations et aux changements d'échelles [Hu62]. Notre base de données d'images est ainsi séparée en deux : deux tiers des images sont conservées pour l'apprentissage, et le tiers restant pour les tests. L'augmentation de données est réalisée

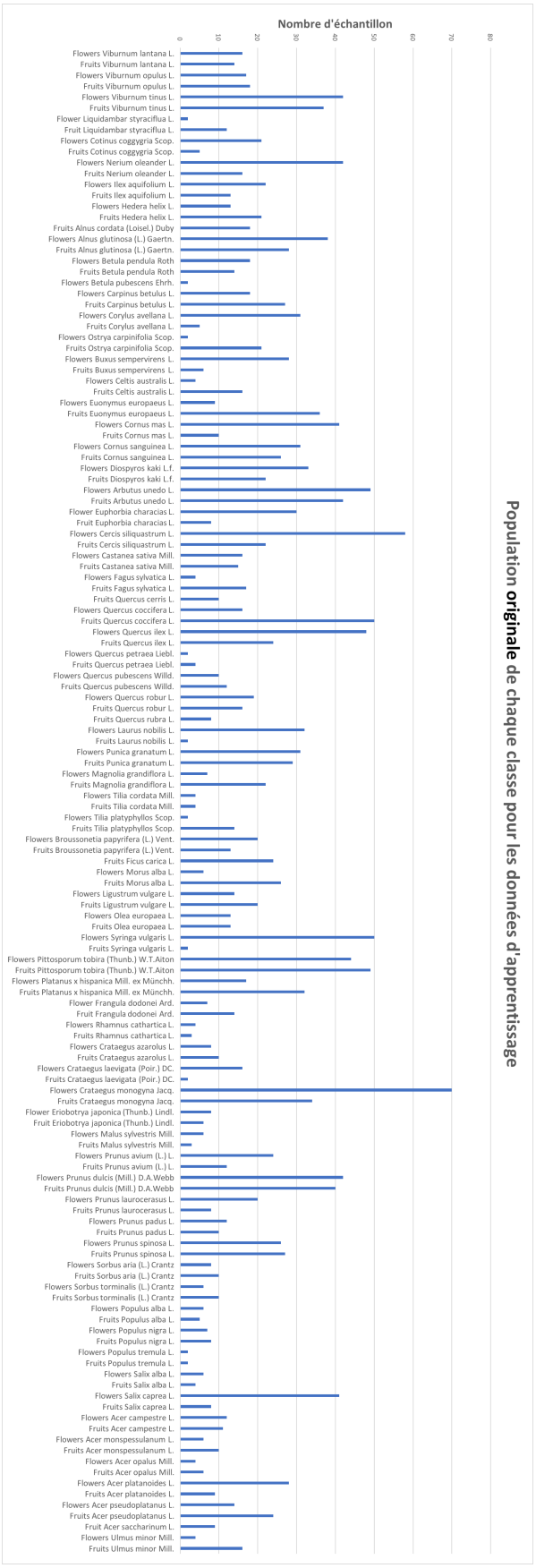


Figure 5.6 – Répartition des images dans les différentes classes.

sur les images appartenant à la base d'apprentissage ; il n'y a pas d'images transformées dans la base de test.

Pour réaliser la transformation de perspective, il est nécessaire de connaître les positions de quatre paires de points entre l'image initiale et l'image transformée. Pour l'image initiale, les quatre points sont les quatre coins de l'image. Pour l'image transformée, chaque point correspond aux coordonnées du point associé de l'image initiale avec un décalage tiré au hasard. Le déplacement du point est tiré parmi des valeurs raisonnables de manière à avoir une légère transformation de perspective qui pourrait être obtenue si l'objet d'intérêt était photographié sous un angle légèrement différent. La transformation de perspective ne peut pas être trop grande sinon cela ne correspondrait pas à la photographie d'un objet 3D. Cette transformation permet de changer la forme de l'objet d'intérêt pour avoir plus de données d'apprentissage avec les moments invariants de Hu (le descripteur de forme). Il n'y a pas de rotation de l'objet réalisée car les moments invariants de Hu sont robustes aux variations d'angle. D'autre part, le changement de perspective permet de modifier légèrement la texture de l'objet d'intérêt. Un exemple d'images augmentées par cette transformation est visible sur la Figure 5.7.

FIGURE 5.7 – Image initiale (à gauche) et deux images augmentées par transformation de perspective.

Par ailleurs, du bruit est ajouté sur les différents canaux colorimétriques de l'image. La moyenne de la gaussienne générant le bruit est égale à zéro tandis que l'écart-type est une valeur aléatoire entre plus ou moins 0,05. Une fois le bruit généré, celui-ci est ajouté indépendamment sur les canaux RGB de l'objet segmenté. Cette transformation permet de générer des données d'apprentissage supplémentaires concernant le descripteur de couleur, ainsi que le descripteur de texture.

En plus d'augmenter le nombre de données, nous pouvons rééquilibrer la base de données en générant plus d'images d'apprentissage pour les classes peu peuplées que pour les classes possédant déjà un nombre suffisant de données. De manière cyclique, les images brutes de chaque classe sont augmentées grâce aux transformations de couleur et de perspective jusqu'à ce que le nombre minimum d'images dans la classe soit de 50.

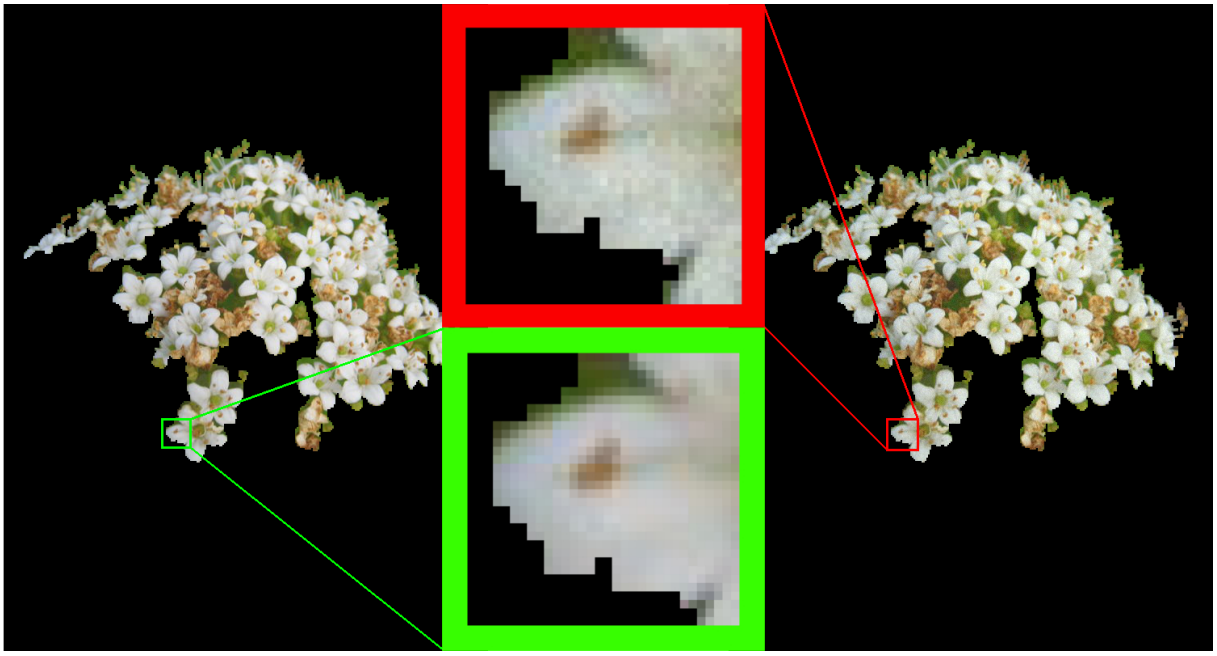


Figure 5.8 – Image initiale (à gauche) et image transformée (à droite) avec la visualisation du bruit ajouté.

5.3.2 Classifieur SVM

En premier lieu, les vecteurs caractéristiques calculés sur les images d'apprentissage sont utilisés pour entraîner un SVM à noyau de séparation RBF, comme pour la classification des écorces (Chapitre 3). Le Tableau 5.1 présente les résultats de classification sur la base de test pour un SVM entraîné avec les données d'apprentissage sans augmentation et avec augmentation de données. Le taux moyen de bonne classification correspond au nombre d'images de test bien classées sur le nombre total d'images de test, tandis que le taux moyen de bonne classification par classe correspond à la moyenne des rapports entre le nombre d'images de test bien classées d'une classe et le nombre d'images de test pour cette classe.

Table 5.1 – Résultats de classification par SVM des fruits et des fleurs avec et sans augmentation de données.

	Taux moyen de bonne classification	Taux moyen de bonne classification par classe
Sans augmentation de données	29,90%	15,20%
Avec augmentation de données	29,55%	21,74%

Nous pouvons voir que le taux de classification moyen n'est pas influencé par l'augmentation de données. Cependant, le taux de classification moyen par classe de fruits-fleurs est bien meilleur après augmentation des données d'apprentissage. Avec les images d'apprentissage brutes, le classifieur SVM tend à sur-prédire certaines classes à cause du

déséquilibre de population entre les classes. Ces résultats sont honorables au vu du grand nombre de classes (136) à reconnaître. Cependant, nous pouvons espérer de meilleur taux, d'autant plus que les fruits et les fleurs sont les organes utilisés par les botanistes pour l'identification des espèces végétales.

5.3.3 Classifieur MLP

Nous proposons par conséquent de tester un autre classifieur pour distinguer les différentes classes. Les vecteurs de caractéristiques des données d'apprentissage vont donc être mis en entrée d'un réseau de neurones de type MLP. Il s'agit plus précisément d'un réseau composé de quatre couches de neurones entièrement connectés. Avoir un réseau avec une seule couche cachée ne permet pas d'obtenir une séparation suffisante entre les classes. Le nombre de neurones sur la première couche correspond à la taille du vecteur de caractéristique, et le nombre de neurones sur la couche de sortie correspond au nombre de classes. Le nombre de neurones sur les deux couches cachées est le même. La variation du nombre de neurones sur les deux couches cachées n'a pas été étudiée pour des raisons de temps, mais en général le nombre de neurones sur les couches cachées doit être compris entre le nombre de neurones de la couche d'entrée et le nombre de neurones sur la couche de sortie. Afin de s'assurer d'avoir un réseau de neurones permettant d'obtenir des résultats positifs, le nombre de neurones sur les couches cachées a été fixé au nombre de neurones sur la couche d'entrée.

Les résultats de classification avec le classifieur MLP sont présentés dans le Tableau 5.2.

Table 5.2 – Résultats de classification par MLP des fruits et des fleurs avec et sans augmentation de données.

	Taux moyen de bonne classification	Taux moyen de bonne classification par classe
Sans augmentation de données	45,19%	36,57%
Avec augmentation de données	47,68 %	40,75%

Nous pouvons noter que, premièrement, avec ou sans augmentation des données d'apprentissage, le MLP donne de bien meilleurs taux de classification que le SVM. De même que pour le classifieur SVM, l'utilisation des images augmentées pour l'apprentissage augmente les résultats de classification. Avec le MLP, les taux de classification sont augmentés en moyenne globale et en moyenne par espèce.

En second lieu, les taux de classification des fruits et des fleurs sont supérieurs aux taux de classification des espèces à partir des images d'écorces. Nous pouvons donc supposer que si nous retournions les espèces à partir de la classe de fruits-fleurs prédite, les taux

de classification des espèces à partir des images de fruits et de fleurs seraient encore supérieurs.

Réduction de l'histogramme de couleurs

Dans cette partie, la méthode de réduction de l'histogramme de couleurs développée dans le Chapitre 3 est appliquée sur la base de données de fruits et de fleurs segmentés. Pour rappel, le calcul de l'histogramme réduit est le suivant :

- L'histogramme de teintes H est calculé sur toutes les images de la base d'apprentissage. Ces histogrammes sont ensuite sommés.
- Le bin de plus faible population est sommé avec son voisin de gauche ou de droite ayant la plus faible population également.
- La fusion des bins est réalisée jusqu'à la taille d'histogramme souhaitée.
- Une matrice de correspondance est aussi créée pour savoir, à chaque fusion de bins, quels sont les bins agrégés. Cette matrice de correspondance permet ensuite de calculer l'histogramme réduit.

Avec l'augmentation des données d'apprentissage, du bruit a été rajouté sur les canaux RGB des images, influençant donc le canal H de l'espace colorimétrique HSV correspondant à la teinte. Le calcul de l'histogramme sommé est donc réalisé sur les images d'apprentissage non augmenté. La Figure 5.9 montre l'histogramme "sommé" pour les images de fruits et de fleurs.

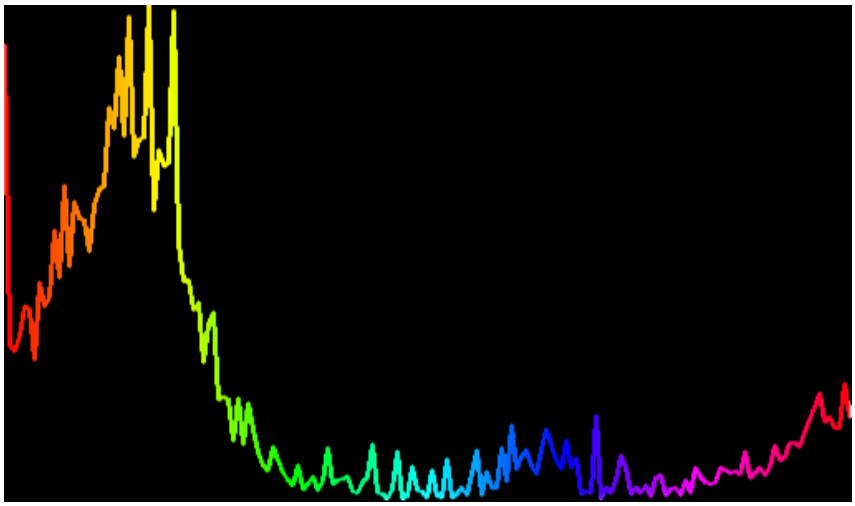


Figure 5.9 – Histogramme "sommé" des canaux H pour les fruits et les fleurs segmentés de la base de données d'apprentissage non augmentée.

L'évolution du taux de classification global et l'évolution du taux de classification en moyenne par classe (fruits-fleurs) sont représentées sur la Figure 5.10 et la Figure 5.11 respectivement.

Contrairement aux écorces, la réduction de l'histogramme de teinte H réduit rapidement le taux de bonne classification. Pour rappel, nous avons limité la réduction de

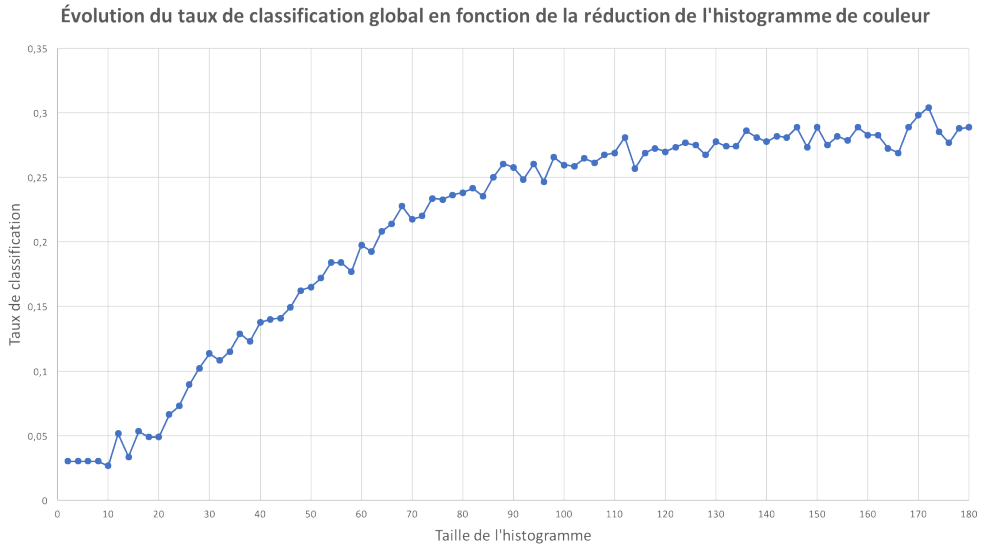


FIGURE 5.10 – Évolution du taux de classification global des fruits et des fleurs classifiés uniquement avec l'histogramme de teintes H en fonction de la taille de celui-ci.

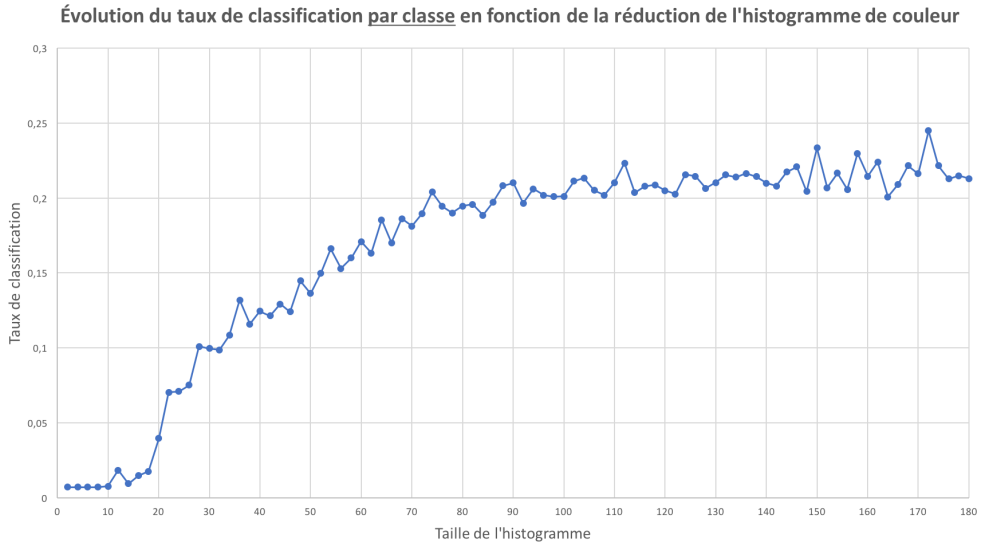


FIGURE 5.11 – Évolution du taux de classification par classe des fruits et des fleurs classifiés uniquement avec l'histogramme de teintes H en fonction de la taille de celui-ci.

l'histogramme de couleurs à 30 bins pour les écorces. Pour la classification des fruits et des fleurs, la couleur joue un rôle beaucoup plus important. Ainsi, nous avons choisi de ne pas effectuer de réduction sur l'histogramme de couleurs.

Prédiction sur les clades botaniques

Le système de classification est entraîné à classifier des fruits et des fleurs indépendamment de leur type, car un utilisateur novice ne peut pas toujours distinguer une inflorescence d'un fruit. De plus, la distinction entre les fruits et les fleurs est ténue car certaines espèces ont leurs fleurs qui se transforment progressivement en fruit, comme nous l'avons vu dans le Chapitre 1 bis. D'autre part, d'un point de vue botanique, les

espèces appartenant à un même genre présentent les mêmes fruits et les espèces appartenant à une même famille ont les mêmes inflorescences. Ceci rajoute de la confusion entre les classes. Par exemple, les fruits de *Quercus ilex* et de *Quercus petraea* peuvent être confondus par le classifieur car ce sont tous les deux des glands.

Nous proposons par conséquent de faire la classification sur les classes de fruits et de fleurs mais de faire la prédiction sur les espèces, les genres et les familles. Si le classifieur a classifié un organe comme "fleur de *Platanus x hispanica*" alors qu'il s'agit d'un fruit de *Platanus x hispanica*, la bonne espèce sera quand même prédite. De même, si le système de classification donne "fruit de *Quercus ilex*" mais qu'il s'agit du fruit de *Quercus petraea*, la prédiction sur le genre (*Quercus*) sera correcte. Nous nous servons pour cela de la correspondance entre les clades de la Figure 5.2 afin de faire les prédictions sur les espèces, les genres et les familles. Les doublons sont supprimés, dans le sens où si le classifieur prédit "fruit de *Quercus ilex*" et " fruit de *Quercus petraea*", nous considérons que le genre *Quercus* n'est prédit qu'une seule fois. Le classifieur utilisé est le MLP, entraîné avec les données d'apprentissage augmentées, qui donne les meilleurs résultats de classification. Les résultats de prédiction sur les clades botaniques sont présentés dans le Tableau 5.3.

Table 5.3 – Résultats de classification par MLP des fruits et des fleurs avec prédiction sur les espèces, les genres et les familles.

	Taux moyen de bonne classification			
	Fruits/Fleurs	Espèces	Genres	Familles
Sans augmentation de données	45,19%	45,79%	50,34%	53,26%
Avec augmentation de données	47,68%	48,02%	51,29%	54,47%

	Taux moyen de bonne classification par classe			
	Fruits/Fleurs	Espèces	Genres	Familles
Sans augmentation de données	36,57%	37,42%	43,02%	45,95%
Avec augmentation de données	40,75%	41,13%	44,08%	46,70%

Au préalable, nous pouvons observer que l'augmentation du nombre de données d'apprentissage augmente les taux de prédiction des espèces, des genres et des familles, puisque ceux-ci sont corrélés à l'augmentation du taux de bonne classification des classes de fruits et de fleurs.

Par ailleurs, comme nous pouvions nous y attendre, plus la prédiction est faite sur un clade élevé (genre ou famille) plus les taux de bonne prédiction augmentent. Nous pouvons remarquer que le taux de bonne classification augmente légèrement entre la classification de la classe fruits-fleur et la prédiction de l'espèce. En effet, il n'y a que certaines espèces qui ont un fruit et une inflorescence très semblables (Figure 5.12).

L'augmentation du taux de bonne prédiction entre la prédiction de l'espèce et du genre montre que certaines espèces ont leurs organes confondus par le classifieur. Notamment,

(a) Fleur de *Betula pendula*. (b) Fruit de *Betula pendula*.

FIGURE 5.12 – Exemple de confusion entre les fruits et les fleurs de l'espèce *Betula pendula*.

il y a confusion entre les fruits des espèces *Acer pseudoplatanus* et *Acer saccharinum* (Figure 5.13).

(a) Fruit de *Acer pseudoplatanus*. (b) Fruit de *Acer saccharinum*.

FIGURE 5.13 – Exemple de confusion entre les fruits des espèces *Acer pseudoplatanus* et *Acer saccharinum*.

Afin d'illustrer la confusion entre les clades botaniques, nous proposons de visualiser les matrices de confusion. En apprentissage supervisé, la matrice de confusion est un outil de mesure de la qualité d'un système de classification. L'un des intérêts de la matrice de confusion est qu'elle indique rapidement si le système parvient à classer correctement ou non les espèces, et elle montre entre quelles classes le classifieur se trompe lors du processus de reconnaissance.

Chaque colonne de la matrice de confusion représente le nombre d'occurrences d'un label prédit, tandis que chaque ligne représente le nombre d'occurrences d'un label vrai. Pour calculer la matrice de confusion, nous avons besoin de données annotées pour tester le classificateur. Pour chaque échantillon de l'ensemble de données d'apprentissage, nous calculons la classe prédite j. Dans la matrice de confusion, nous ajoutons 1 dans la cellule $M_{i,j}$ où i est le label vrai et j le label prédit. Une fois que toutes les données

de test sont classifiées, la matrice de confusion est normalisée sur les lignes. Ainsi, les résultats de la matrice de confusion sont moins influencés par le nombre d'échantillons de test dans chaque classe. Au final, les valeurs $M_{i,j}$ appartiennent à l'intervalle [0,1] (voir Algorithme 1).

Algorithm 1 Algorithme de calcul de la matrice de confusion.

```
int nbTestSample ← nombre d'images de test
for (int a = 0 ; a < nbTestSample ; a++) do
  if j == i then
    M_{i,j}++
  end if
end for
M_{i,j} ← M_{i,j} / ∑ nbTestSample ∈ i        //normalisation sur les lignes
```

La matrice de confusion d'un classifieur efficace présente des valeurs élevées sur la diagonale et des valeurs basses dans le reste de la matrice. Il est possible de visualiser une matrice de confusion en utilisant une échelle de couleurs allant du blanc au vert, blanc pour les valeurs basses et vert pour les valeurs élevées. Ainsi, un classifieur parfait génère une matrice de confusion avec seulement des cellules vertes sur la diagonale.

La Figure 5.14 présente la matrice de confusion sur la prédiction des espèces.

Cette matrice de confusion témoigne que le classifieur utilisé permet globalement de bien classifier les différentes classes de fruits et de fleurs, et ainsi de correctement prédire les espèces d'arbres. La Figure 5.15 présente un détail de cette matrice centré sur les espèces appartenant au genre *Quercus* (les chênes).

Ce détail montre qu'il y a de la confusion entre les différentes espèces de chênes. Par exemple, *Quercus coccifera* est confondu, entre autres, avec *Quercus ilex* et *Quercus pubescens*. Pour *Quercus rubra*, les organes ont été mal classifiés par le MLP et confondus avec des espèces n'appartenant pas au genre *Quercus*. Pour la prédiction sur les genres, toutes ces espèces sont rassemblées sous la même étiquette. Ainsi, la Figure 5.16 montre la matrice de confusion pour la prédiction du genre.

Enfin, nous pouvons noter une augmentation du taux de bonne identification entre la prédiction sur le genre et sur la famille. La prédiction sur le dernier clade permet de rendre compte de la confusion qu'il peut y avoir au niveau de la famille botanique, particulièrement avec les genres appartenant à la famille des *Betulaceae*, comme le montre le détail de la matrice de confusion sur les genres 5.17, dont toutes les inflorescences sont des chatons 5.18.

Pour finir, la Figure 5.19 présente la matrice de confusion obtenue avec la prédiction sur les familles.

Si nous nous replaçons dans le cadre de l'utilisation de l'application *Folia* pour l'aide à l'identification d'espèces d'arbres, nous pouvons proposer une liste de $N = 10$ clades botaniques les plus probables auquel un fruit ou une fleur peut appartenir. Les pourcentages

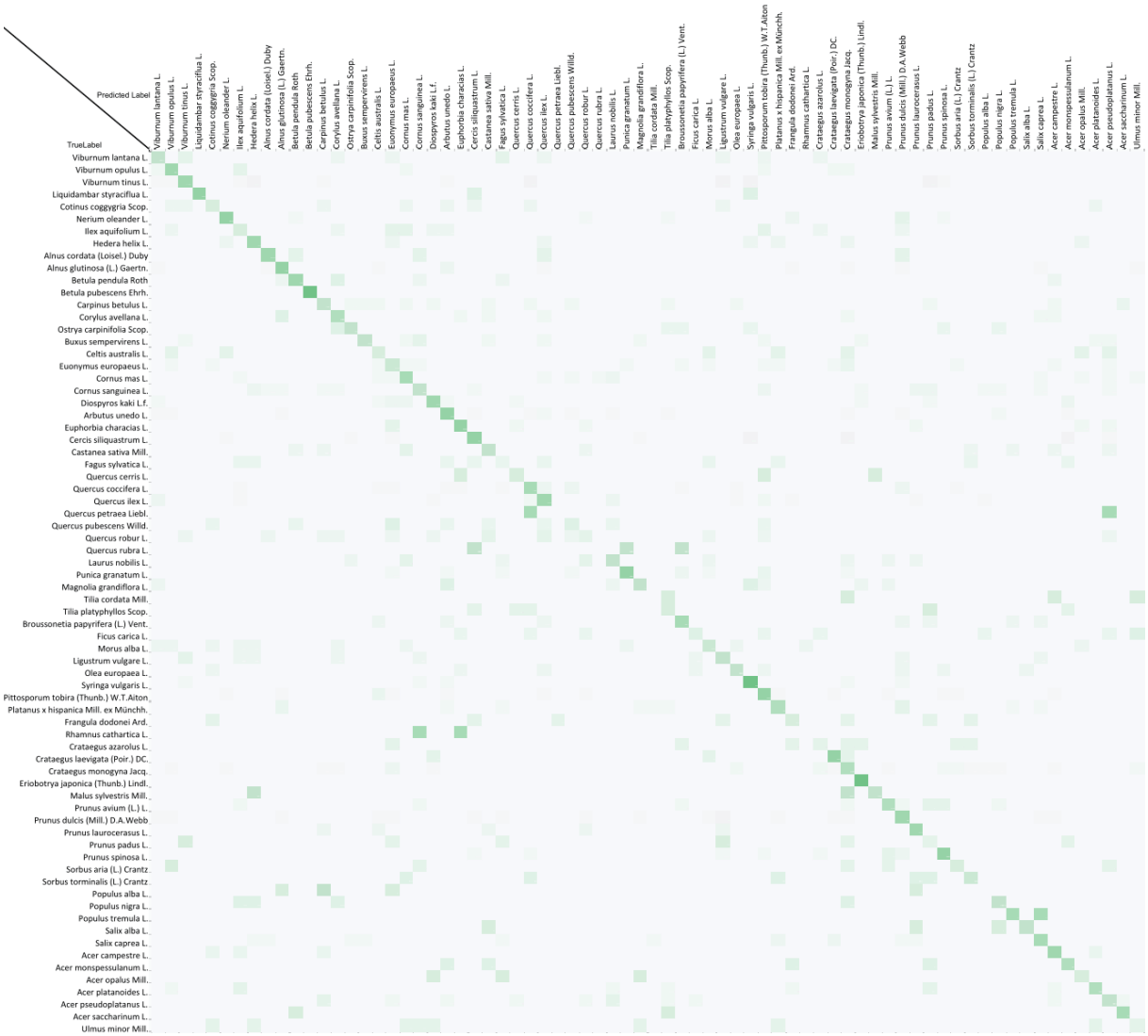


FIGURE 5.14 – Matrice de confusion issue de la prédiction des espèces.

TrueLabel \ Predicted Label	Quercus cerris L.	Quercus coccifera L.	Quercus ilex L.	Quercus petraea Liebl.	Quercus pubescens Willd.	Quercus robur L.	Quercus rubra L.
Quercus cerris L.	0,25	0	0	0	0	0	0
Quercus coccifera L.	0	0,55	0,06	0	0,06	0	0
Quercus ilex L.	0	0,06	0,56	0	0	0	0
Quercus petraea Liebl.	0	0,5	0	0	0	0	0
Quercus pubescens Willd.	0	0	0,09	0	0,18	0	0
Quercus robur L.	0	0	0,13	0	0,06	0,13	0
Quercus rubra L.	0	0	0	0	0	0	0

FIGURE 5.15 – Détail de la matrice de confusion sur la prédiction des espèces (Figure 5.14) pour les espèces appartenant au genre *Quercus*.

de chance d'obtenir le bon clade dans une liste pour N allant de 1 à 10 sont présentés sur la Figure 5.20. Dans le cas de l'utilisation de *Folia* où N est égal à 10, il y a 88,75%

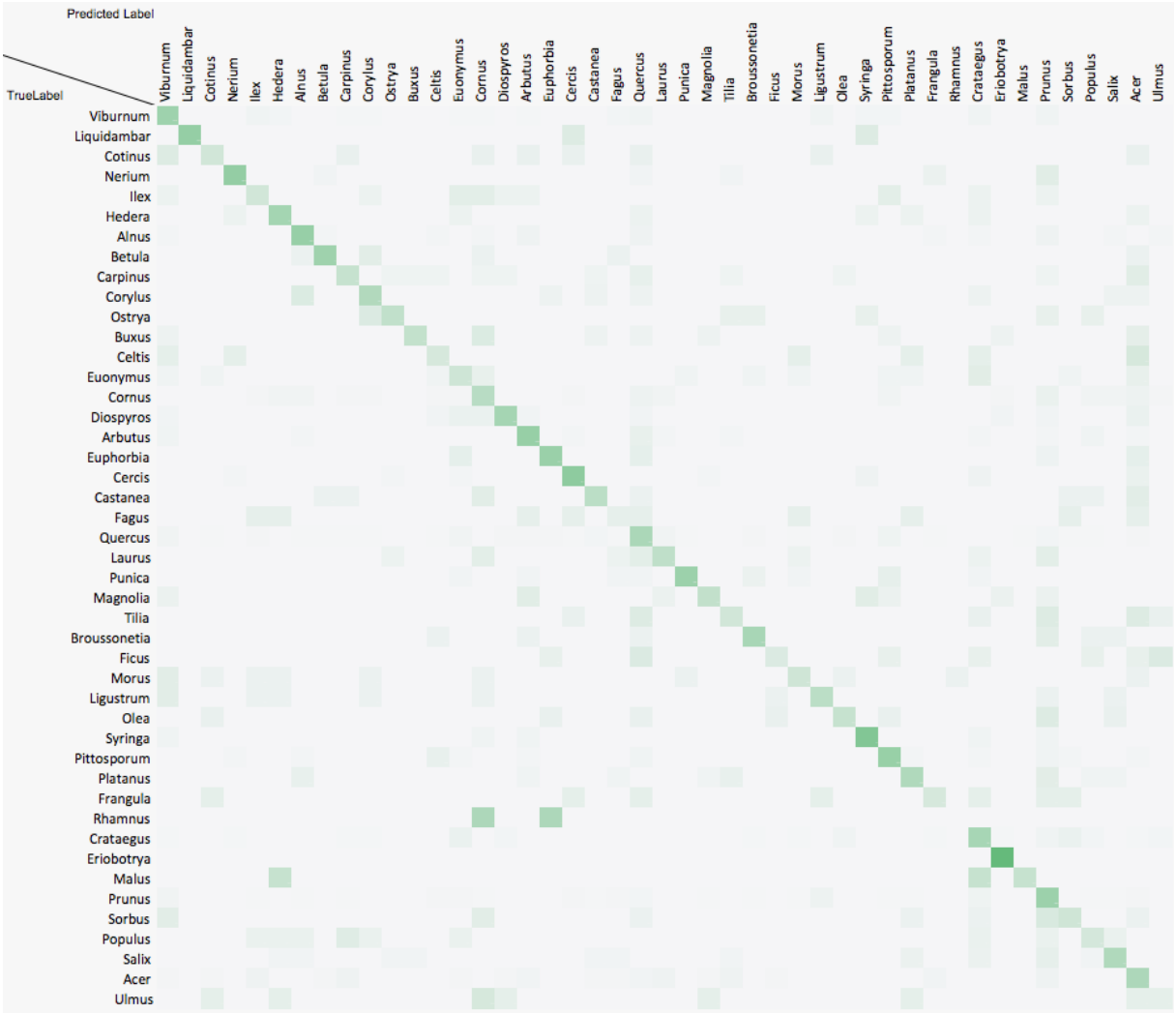


Figure 5.16 – Matrice de confusion issue de la prédiction des genres.

Predicted label / True label	Alnus	Betula	Carpinus	Corylus	Ostrya
Alnus	0,66	0,02	0	0	0
Betula	0,07	0,6	0	0,13	0
Carpinus	0	0	0,33	0	0,05
Corylus	0,18	0	0	0,47	0
Ostrya	0	0	0	0,18	0,36

Figure 5.17 – Détail de la matrice de confusion sur la prédiction des genres (Figure 5.16) pour les genres appartenant à la famille des *Betulaceae*.

de chance d'avoir la bonne famille dans la liste proposée, 82,99% de chance d'obtenir le bon genre, 76,37% de chance d'obtenir la bonne espèce, et 73,63% de chance d'obtenir la bonne identification entre le fruit et la fleur.

(a) *Alnus*. (b) *Betula*. (c) *Carpinus*. (d) *Corylus*.

FIGURE 5.18 – Fleurs de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae*.

FIGURE 5.19 – Matrice de confusion issue de la prédiction des familles.

5.4 Comparaison avec des méthodes de Deep Learning

Les descripteurs mis en place sont très simples et ne présentent pas une originalité de recherche. Ils montrent cependant leur efficacité pour la classification sur un grand nombre

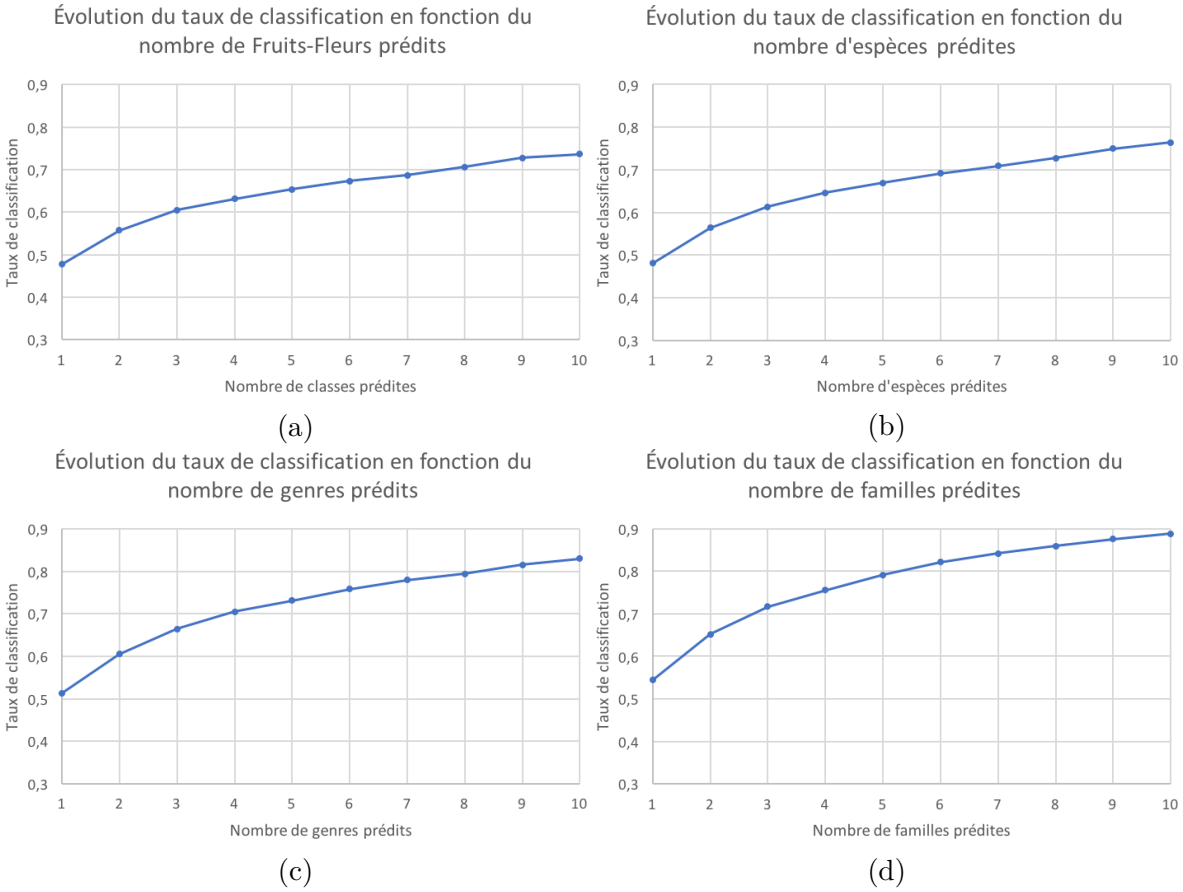


Figure 5.20 – Évolution du taux de classification pour la classification des classes de fruits-fleurs a), et pour la prédiction sur les espèces b), les genres c), les familles d).

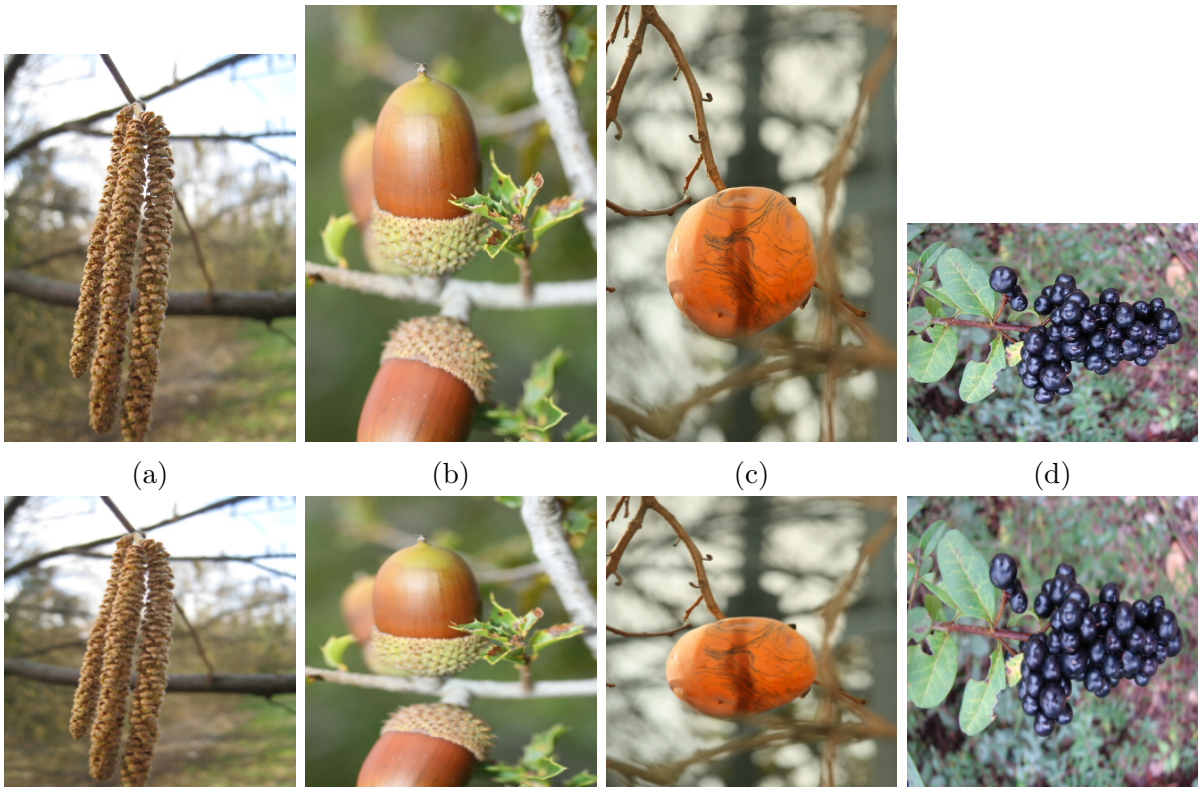
de classes de fruits et de fleurs. Cette partie présente des résultats de classification pour notre base de données en utilisant des méthodes de Deep Learning qui ont fait leurs preuves dans d'autres contextes [GBJ16]. De même que pour les écorces, le nombre de données n'est pas avantageux pour ce type de méthodes. Les réseaux de neurones à convolutions que nous testons sont AlexNet et GoogLeNet entraînés sur notre base de données.

5.4.1 Images non segmentées

Premièrement, les performances de classification des CNN sont évaluées sur les images non segmentées. En effet, un des avantages de l'utilisation des CNN est la possibilité de fournir les images brutes en entrée du réseau sans nécessité de segmentation de l'objet d'intérêt.

Les images utilisées sont les mêmes que celles utilisées pour obtenir les résultats précédents mais avant leur étape de segmentation. Il y a donc 136 classes de fruits et de fleurs, correspondant à 72 espèces d'arbres. Il n'y a pas de distinction entre les fleurs et les fruits. Pour réaliser l'apprentissage, deux tiers des images sont utilisés, dont 10% servent à la validation. Le dernier tiers est utilisé pour les tests. Étant donné que les CNN prennent

des images de taille fixe en entrée, les images sont redimensionnées pour que la largeur et la hauteur correspondent à l'entrée du réseau. Cette transformation (Figure 5.21) n'est peut-être pas la plus appropriée puisque la fleur ou le fruit pourra voir sa forme altérée. De manière optimale, il faudrait que l'image soit retaillée de manière à avoir la bonne taille et que l'objet d'intérêt soit au centre, sans que certaines parties de l'inflorescence ou du fruit ne soit rognées. Cependant cette solution n'a pas été mise en place faute de temps puisqu'il aurait fallu faire cela "à la main".



(a) (b) (c) (d)

(e) (f) (g) (h)

FIGURE 5.21 – Exemple d'images originales a), b), c) d), et leur image retaillée associée e), f), g), h) pour mettre en entrée des CNN.

Le Tableau 5.4 présente les résultats de classification avec les réseaux de neurones à convolutions AlexNet et GoogLeNet.

TABLE 5.4 – Performances des CNN AlexNet et GoogLeNet sur des images non segmentées.

	Top 1	Top 5
AlexNet	33,48%	57,74%
GoogLeNet	27,15%	54,03%
Notre méthode (sans augmentation de données)	45,18%	63,75%

D'après ces taux de bonne classification obtenus sur les images brutes, nous pouvons remarquer que les descripteurs qui ont été choisis permettent de donner de meilleurs résultats. Ceci peut s'expliquer par le fait que les CNN doivent apprendre par eux-mêmes

les caractéristiques pertinentes à extraire sur les images brutes et non directement sur l'objet d'intérêt qui n'a pas été préalablement extrait de l'image. Ces caractéristiques sont d'autant plus difficiles à extraire du fait du grand nombre de classes et de la faible population de celles-ci. Ainsi, le CNN GoogLeNet qui possède plus de couches de convolution donne de moins bons résultats car il n'a vraisemblablement pas assez de données d'apprentissage pour le calcul des filtres. Par ailleurs, l'objet d'intérêt peut se trouver en plusieurs exemplaires sur l'image brute.

5.4.2 Images segmentées

Dans un second temps, les réseaux AlexNet et GoogLeNet ont été entraînés sur les images segmentées qui ont été utilisées pour l'extraction des descripteurs de formes, de couleurs et de textures. De même, il y a 136 classes de fruits ou de fleurs correspondant à 72 espèces. L'apprentissage des réseaux est réalisé sur deux tiers des images de la base de données, dont 10% pour la validation, et les tests sont réalisés sur le reste de la base. De la même manière que pour les images non segmentées, les images sont redimensionnées pour avoir la taille requise en entrée des réseaux. Les CNN ont été entraînés avec les données d'apprentissage brutes et les données augmentées qui ont été présentées dans la partie 5.3.1. Le Tableau 5.5 présente les résultats de classification des réseaux AlexNet et GoogLeNet entraînés avec, et sans, augmentation de données.

Table 5.5 – Performances des CNN AlexNet et GoogLeNet sur des images segmentées avec, et sans, augmentation de données.

	Sans augmentation de données		**Avec augmentation de données**	
	Top 1	Top 5	Top 1	Top 5
AlexNet	12,59%	26,38%	13,19%	28,10%
GoogLeNet	14,66%	30,26%	16,12%	33,53%
Notre méthode	45,18%	63,75%	47,68%	65,38%

Par rapport à l'utilisation des CNN entraînés sur les images brutes, nous pouvons voir qu'entraîner des réseaux de neurones convolutionnels sur des images segmentées n'améliore pas les résultats de classification, mais au contraire les fait fortement baisser. Bien que l'augmentation du nombre d'images segmentées pour l'apprentissage augmente les taux des CNN, ceux-ci restent bien inférieurs à ceux que nous obtenons avec notre vecteur de caractéristiques. Par contre, à l'opposé de l'utilisation des images brutes pour entraîner les CNN, GoogLeNet obtient de meilleurs taux de classification que AlexNet. Ici le nombre de données d'apprentissage n'est pas vraiment à mettre en cause. L'origine de la baisse des performances des CNN sur les images segmentées pourrait venir du fait que des informations utiles à l'identification des classes pourraient se trouver dans le fond des images, tel que le feuillage.

5.5 Conclusion

Dans ce chapitre, nous avons privilégié certains descripteurs afin de créer un vecteur de caractéristiques permettant de différencier les différentes classes de fleurs et de fruits de 72 espèces d'arbre. Ces vecteurs ont été extraits sur des organes segmentés. Malgré le fait que les inflorescences et les fruits ne sont pas séparés, les résultats de classification sont convenables.

Pour pallier le problème du manque de données, nous avons proposé une méthode pour générer de nouvelles images d'apprentissage. Deux classifieurs ont été testés pour identifier les différents classes, et le MLP a montré sa supériorité. L'augmentation de données a aussi prouvé son efficacité en augmentant les taux de bonnes classifications. D'autre part, l'utilisation de la segmentation et de notre vecteur de caractéristiques extrait sur les organes segmentés permet une meilleure identification de la classe par rapport à l'utilisation des réseaux de neurones à convolutions, et ce avec ou sans augmentation de données.

Enfin, en prenant en compte la nature botanique des objets, nous avons proposé d'effectuer une prédiction sur les différents clades botaniques (espèce, genre, famille) à partir des prédictions du classifieur sur les classes de fruits et de fleurs. Les prédictions d'appartenance au genre ou à la famille issues des fleurs ou des fruits seront une source d'information pour l'identification des espèces à partir des images de feuille ou d'écorce. De même que pour les clefs botaniques, une restriction des espèces probables pourra être réalisée à partir d'une image de fleur, avant que la prédiction de l'espèce ne soit affinée par une photo de feuille.

Chapitre 6

Stratégies de fusion des données pour la classification multi-organes d'espèces végétales

Dans ce chapitre, nous allons nous intéresser à la fusion des données extraites des photos de feuilles dans l'application *Folia* et des différents vecteurs de caractéristiques d'écorces et de fruits ou de fleurs dont les compositions ont été présentées dans les chapitres précédents. Nous allons particulièrement nous intéresser à la fusion des données de feuille et d'écorce. En effet, ces deux organes sont observables la plus grande majorité de l'année, contrairement aux fruits et aux fleurs. Dans un premier temps, la composition du vecteur de feuille est présentée. Deux méthodes de combinaisons basiques sont ensuite proposées. La contribution principale de ce chapitre est la prise en compte de la confusion entre les espèces grâce aux matrices de confusion. La dernière partie de ce chapitre propose différentes méthodes de prise en compte des informations botaniques de clade obtenues grâce aux fruits et aux fleurs. Il s'agit finalement d'étudier l'influence de la prise en compte des différents organes lors de l'identification d'espèces d'arbres, et plus particulièrement de s'assurer que celle-ci n'engendre pas une baisse du taux de classification par rapport à la classification des feuilles seules.

6.1 Identification des feuilles dans l'application *Folia*

Le descripteur de feuilles ainsi que le classifieur utilisés dans l'application *Folia* ont été développés par Guillaume Cerutti *et al.* [Cer+13b] dans le cadre du projet ANR ReVeS. Les détails et les choix réalisés pour la construction de ce vecteur et de ce classifieur sont détaillés dans sa thèse intitulée "Segmentation et interprétation d'images naturelles pour l'identification de feuilles d'arbres sur smartphone" [Cer13b]. Un résumé des principaux éléments est cependant présenté dans cette partie.

6.1.1 Descripteurs de feuilles

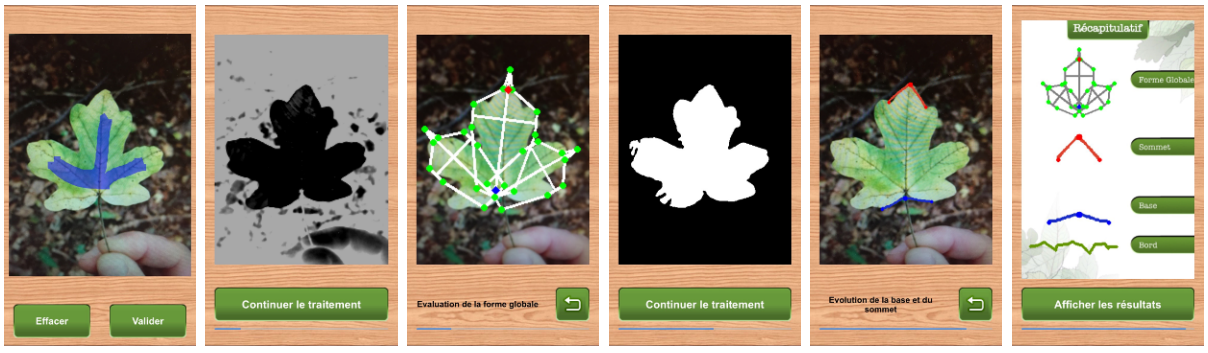


FIGURE 6.1 – Les différentes étapes de traitement pour l'identification de la feuille par l'application *Folia*.

Dans un premier temps, la feuille est segmentée grâce à un modèle paramétrique de forme déformable (troisième image de la Figure 6.1), qui s'applique sur tout type de forme foliaire. Ce modèle déformable est appliqué sur une carte de distance de couleur de

la feuille (deuxième image de la Figure 6.1) calculée à partir des couleurs sélectionnées par une trace de l'utilisateur sur la feuille (trace bleue sur la première image de la Figure 6.1). C'est ensuite un algorithme de contour actif, initialisé et contraint par le modèle paramétrique, qui va venir segmenter précisément la feuille d'intérêt (quatrième image de la Figure 6.1). Cette méthode de segmentation a été comparée à celles de la littérature appliquées aux images de feuilles dans un environnement naturel et son efficacité a été démontrée [Gra+15].

Les descripteurs développés pour la caractérisation des feuilles sont basés sur les informations morphologiques extraites par les botanistes, c'est-à-dire la forme générale de la feuille, la forme de l'apex, la forme de la base, et la marge. Ainsi, les paramètres du modèle déformable (**Polygone**) sont utilisés, de même que le nombre de lobes (**Lobes**) dans le vecteur descripteur. En effet, ces paramètres décrivent la forme générale de la feuille. De plus, l'apex et la base de la feuille sont pris en compte. Des courbes de Bézier y sont ajustées et les paramètres de ces courbes sont considérés comme des descripteurs (**Base** et **Apex**). La cinquième image de la Figure 6.1 montre un exemple de telles courbes (rouge pour l'apex et bleu pour la base). Enfin, la dentition de la feuille est également décrite. Pour ce faire, une méthode basée sur le calcul de la courbure de la marge de la feuille associée à une description à l'aide d'un vocabulaire est utilisée (**MotBord**). Le vecteur caractéristique de feuille, de dimension 28, est finalement le suivant :

$$\mathbf{F}_l = \{\mathbf{Polygone}, \mathbf{Lobes}, \mathbf{Base}, \mathbf{Apex}, \mathbf{MotBord}\} \tag{6.1}$$

6.1.2 Classifieur de feuilles

L'application *Folia* utilise une classification basée sur le voisin le plus proche dans laquelle le classifieur construit un modèle de chaque espèce $\mathbf{S}$ comme une distribution gaussienne multidimensionnelle de ses vecteurs de caractéristiques, et estime la distance d'un nouvel échantillon représenté par son vecteur de caractéristiques $\mathbf{F}_l$ à tous les modèles d'espèces. Pour éviter que les espèces à variabilité élevée absorbent des classes moins variables, la distance $d_{GNN}(\mathbf{F}_l, \mathbf{S})$ est définie comme la distance euclidienne à la surface de l'hyperellipsoïde définie par son centroïde et la matrice de covariance de l'espèce. Le résultat de cette classification est l'espèce qui correspond à la distance minimale. La même approche de classification peut être utilisée pour les caractéristiques de l'écorce $\mathbf{F}_b$, des fruits $\mathbf{F}_{fr}$ ou des fleurs $\mathbf{F}_{fl}$ et les vecteurs de caractéristiques concaténées.

6.2 Base de données

Nous souhaitons dans un premier temps faire la fusion entre les données de feuilles et les données d'écorces. L'application *Folia* permet d'identifier 120 espèces d'arbres à partir

de leurs feuilles. Pour l'identification des écorces, nous avons une base de 101 espèces. Ces deux bases sont issues de la base PlantCLEF. Cependant, certaines espèces de cette base ont des photos de feuilles mais pas d'écorces, de fruits, de fleurs, ou *vice versa*. De ce fait, le nombre d'espèces que nous pouvons considérer est réduit aux 72 classes utilisées dans le chapitre précédent traitant de la classification des fruits et des fleurs.

6.3 Méthodes de combinaisons basiques

Les deux méthodes qui vont être présentées dans les parties 6.3.1 et 6.3.2 permettent de fusionner les informations issues des vecteurs caractéristiques des feuilles et des écorces sans prendre en compte d'information *a priori*. Pour rappel, nous dénotons $\mathbf{F}_l$ le vecteur descripteur de feuilles et $\mathbf{F}_b$ le vecteur descripteur d'écorce.

6.3.1 Concaténation des vecteurs

L'un des systèmes de fusion les plus simples est de concaténer directement le vecteur caractéristique de la feuille et le vecteur caractéristique de l'écorce, et de donner ce vecteur concaténé à un classifieur. Nous utilisons deux classifieurs : le classifieur de *Folia*, détaillé dans la partie 6.1.2, et le classifieur SVM un contre tous, présenté dans le Chapitre 3.

Les Figures 6.2 et 6.3 montrent les taux de reconnaissance du vecteur caractéristique de la feuille seulement $\mathbf{F}_l$ (courbes vertes), du vecteur caractéristique de l'écorce seulement $\mathbf{F}_b$ (courbes marron) et de la concaténation des vecteurs caractéristiques de feuille et d'écorce $\mathbf{F}_{l,b}$ (courbes bleues), respectivement avec les deux systèmes de classification différents (*Folia* et SVM).

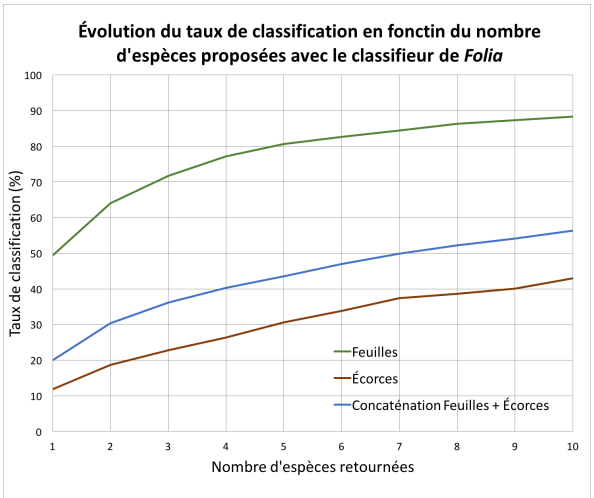


FIGURE 6.2 – Taux de reconnaissance des vecteurs $\mathbf{F}_l$, $\mathbf{F}_b$ et $\mathbf{F}_{l,b}$ avec le classifieur de *Folia*.

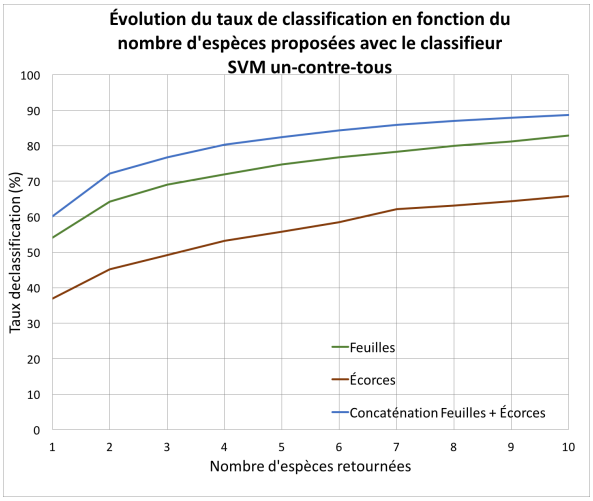


FIGURE 6.3 – Taux de reconnaissance des vecteurs $\mathbf{F}_l$, $\mathbf{F}_b$ et $\mathbf{F}_{l,b}$ avec le classifieur SVM un-contre-tous.

La Figure 6.2 présente les taux de classification avec le classifieur utilisé dans l'application *Folia*. Celui-ci donne de bons résultats pour la classification des feuilles avec 49,5% de bonne prédiction, et jusqu'à 88,4% de chance d'avoir la bonne espèce dans une liste de 10 espèces. Cependant les taux de classification pour les écorces sont très faibles. En effet, le classifieur de *Folia* a été établi pour des vecteurs de très petites tailles, ce qui n'est pas exactement le cas pour notre vecteur caractéristique d'écorces. La concaténation des deux vecteurs ne donne pas non plus de bons résultats de classification, avec seulement 20,0% de chance de retrouver la bonne espèce, ce qui est inférieur à la classification de la feuille uniquement.

Les taux de classification obtenus avec le classifieur SVM un contre tous sont présentés sur la Figure 6.3. Premièrement, nous pouvons noter que les taux de classification des écorces (courbe marron) sont supérieurs au taux du Chapitre 3 car le nombre d'espèces de la base de données est plus faible (72 espèces ici contre 101 dans le chapitre consacré à la reconnaissance des écorces). Dans un second temps, nous pouvons voir que ce classifieur, qui a été développé pour l'identification des écorces, donne de bien meilleurs résultats pour cet organe que le classifieur de *Folia*. En effet, le taux de classification commence à 36,9% (au lieu de 11,8%) et va jusqu'à 65,9% de chances de trouver la bonne espèce sur une liste de dix espèces en utilisant une image d'écorce (au lieu de 43,0%). Pour l'identification de la feuille, le classifieur basé SVM est plus susceptible de trouver la bonne espèce avec un taux de classification de 54,1%. Mais à partir de 3 espèces proposées, c'est le classifieur de *Folia* qui fournit de meilleurs résultats : 71,7% (Figure 6.2) au lieu de 69,1% pour le classifieur SVM.

Enfin, nous pouvons examiner le taux de classification pour le vecteur concaténé (courbe bleue) obtenu avec le système de classification SVM (Figure 6.3). Avec cette méthode de classification, le taux de bonne classification est de 60,3% si une seule espèce est proposée à l'utilisateur, 82,4% pour les 5 premières espèces les plus probables, et jusqu'à 88,6% de chance de trouver l'espèce correcte dans une liste de dix espèces proposées. Nous pouvons donc voir que la prise en compte de l'écorce augmente le taux d'identification de l'espèce d'arbre par rapport au cas où uniquement la feuille est considérée.

Nous pouvons cependant espérer une amélioration de ces résultats par une méthode de combinaison de données un peu moins basique.

6.3.2 Combinaison *a posteriori*

Nous pouvons émettre l'hypothèse que la feuille et l'écorce n'ont pas la même influence sur la reconnaissance d'une espèce d'arbre. En effet, d'après nos résultats précédents, nous pouvons voir que nous avons plus de chances de récupérer la bonne espèce à partir d'une photo de la feuille que d'une photo de l'écorce. C'est pourquoi nous avons mis en place une méthode de combinaison *a posteriori*. Autrement dit, nous faisons les classifications

indépendamment pour l'écorce et pour la feuille, puis nous combinons les résultats de ces deux classifications. Avec cette approche, nous avons la possibilité d'équilibrer l'influence de la feuille ou de l'écorce dans le processus d'identification. Notre méthode de combinaison *a posteriori* utilise toutes les réponses des classifieurs sous forme de confiances (entre 0 et 1) associées à chaque espèce. Ensuite, la combinaison est réalisée sous la forme d'une fusion multiplicative des valeurs de confiance pour chaque espèce :

$$\overset{\otimes}{c}(\mathbf{F}_b, \mathbf{F}_l, \mathbf{S}) = \Big(c(\mathbf{F}_b, \mathbf{S})\Big)^{\frac{1}{\omega_b}} \Big(c(\mathbf{F}_l, \mathbf{S})\Big)^{\frac{1}{\omega_l}} \tag{6.2}$$

avec ω_l et ω_b les poids associés respectivement à la feuille et à l'écorce.

La confiance résultante $\overset{\otimes}{c}$ est utilisée pour ordonner les espèces, et le résultat de classification est la liste des espèces correspondant à la confiance combinée de la plus élevée à la moins élevée. Les résultats sont présentés sur la Figure 6.4.

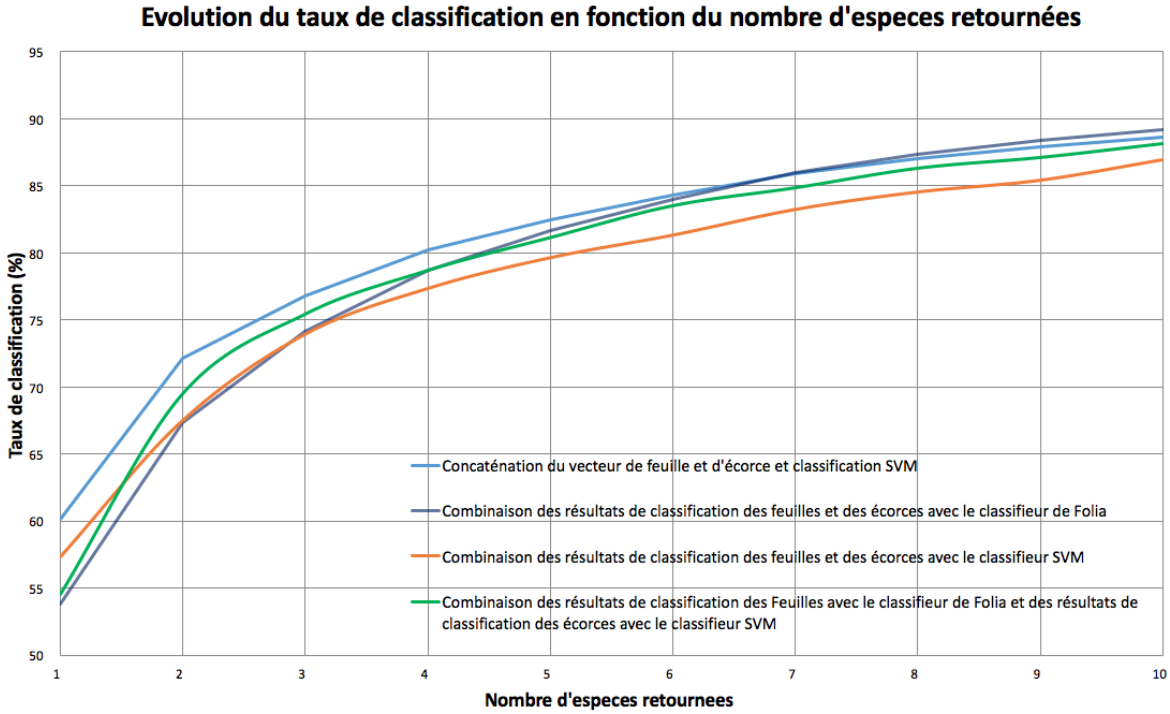


FIGURE 6.4 – Combinaison *A posteriori* du vecteur de feuille $\mathbf{F}_l$ et du vecteur d'écorce $\mathbf{F}_b$.

Dans tous les cas, nous pouvons remarquer que la combinaison *a posteriori* augmente le taux de reconnaissance de l'arbre par rapport à l'écorce seule ou à la feuille seule quel que soit le système de classification (SVM ou *Folia*). Le meilleur taux obtenu avec une combinaison *a posteriori* en top 1 est obtenu avec les résultats de classification des feuilles et des écorces réalisés avec le classifieur SVM (courbe orange). De la partie précédente, nous avons vu que nous avons obtenu le meilleur taux de classification pour la reconnais-

sance des feuilles seulement avec le classifieur Folia et avec le classifieur basé sur SVM pour l'écorce seulement. Ainsi, dès le top 2, il faut privilégier la combinaison *a posteriori* entre ces deux systèmes de classification (courbe verte). Cependant, à partir d'une liste de 5 espèces proposées à l'utilisateur, il faut privilégier la combinaison *a posteriori* des résultats de classification des feuilles et des écorces avec le classifieur SVM (courbe bleu foncé) pour avoir le plus de chance d'avoir la bonne espèce dans la liste. Pour comparaison, nous affichons les résultats des vecteurs de caractéristiques concaténés des feuilles et de l'écorce avec le classifieur SVM qui a donné le meilleur taux de classification dans la partie précédente (courbe bleu clair). Nous pouvons remarquer que c'est cette simple concaténation des vecteurs de feuilles et d'écorce classés avec SVM qui donne le meilleur taux de classification jusqu'au top 6 (courbe bleu clair).

En conclusion de ce paragraphe, nous pouvons dire que la simple concaténation des vecteurs descriptifs et l'utilisation de SVM pour les classifier semblent pour le moment la solution la plus efficace tout en étant la plus simple à mettre en œuvre. Cependant, même si la prise en compte *a posteriori* des deux vecteurs améliore les taux de reconnaissance des espèces d'arbres, cette dernière reste modeste et nous pouvons imaginer d'en obtenir de meilleurs en développant des méthodes plus évoluées. En particulier, l'idée est de prendre en compte les matrices de confusion pour raffiner la contribution de la feuille et de l'écorce dans la reconnaissance

6.4 Prise en compte de la confusion entre les espèces

Une des contraintes de l'identification des végétaux est la similarité qu'il y a entre les différentes espèces. En effet, la similarité extra-classe est notamment importante entre les différentes classes d'écorces comme nous avons pu le voir dans le Chapitre 3, ou même entre les feuilles. De même, il y a une très grande similarité extra-classe entre les différentes espèces pour les fleurs et les fruits. Cette similarité induit le classifieur en erreurs. Afin de prendre en compte cette particularité, nous proposons de considérer les matrices de confusion. Pour rappel, la matrice de confusion permet de montrer entre quelles classes le classifieur se trompe lors du processus de classification. La matrice de confusion est calculée selon le même algorithme que celui utilisé dans le chapitre traitant de la reconnaissance des fleurs et des fruits (Chapitre 5).

6.4.1 Matrices de confusion issues de la classification des feuilles et la classification des écorces

Les Figures 6.5 et 6.6 présentent les matrices obtenues respectivement pour la classification uniquement des feuilles avec le classifieur de *Folia* et la classification uniquement des écorces avec le classifieur SVM un contre tous. Rappelons que ce sont ces configura-

tions qui donnent les meilleurs taux de classification (partie 6.3).

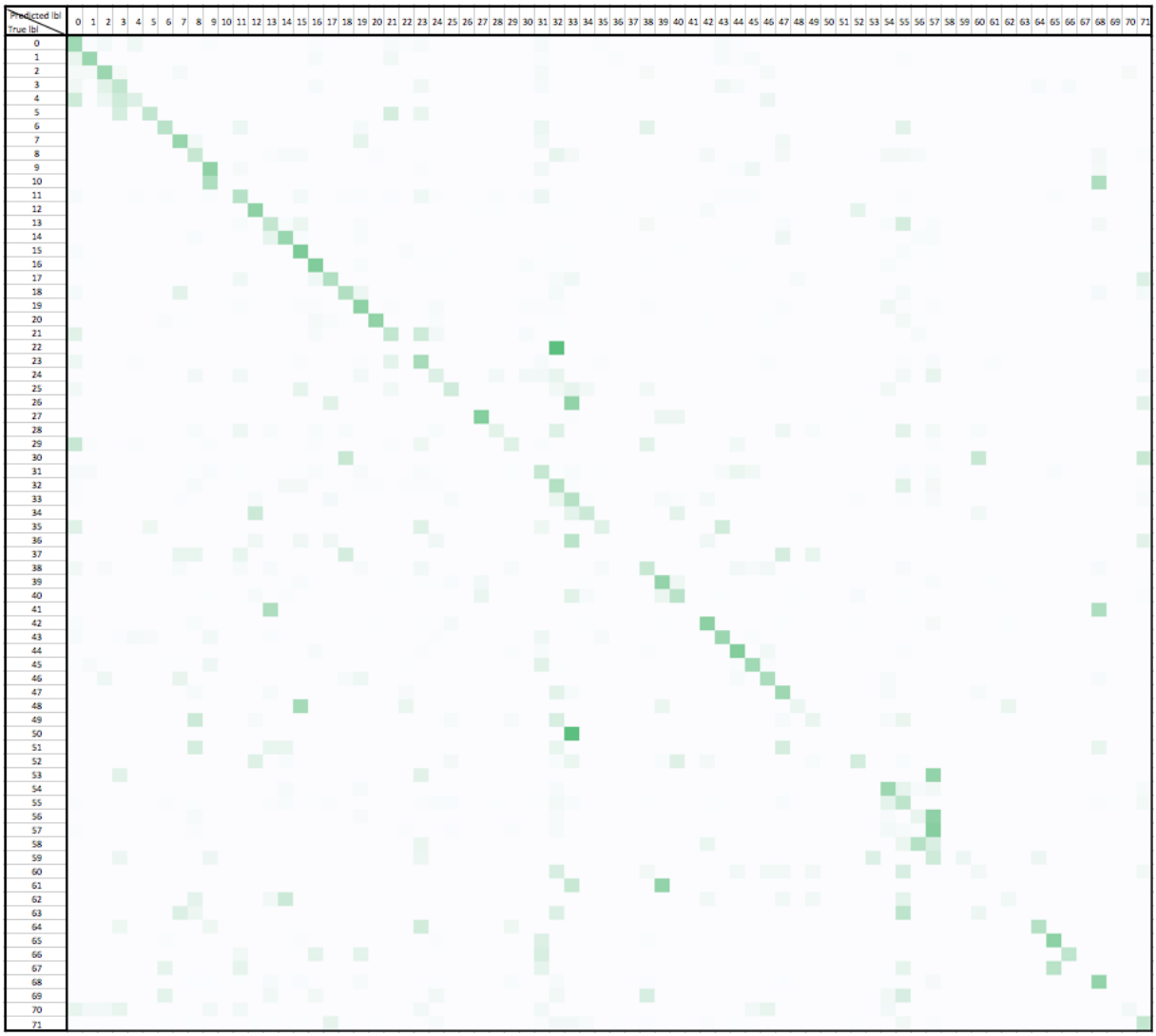


FIGURE 6.5 – Matrice de confusion issue de la classification des feuilles grâce au classifieur de *Folia*.

En comparant ces matrices, nous pouvons remarquer que le classifieur de *Folia* utilisé pour les feuilles est en fait plus fiable que le classifieur SVM utilisé pour les écorces. Il se trouve en effet que la diagonale est plus contrastée pour cette dernière. Sur la Figure 6.6, nous pouvons clairement observer que les écorces sont souvent confondues entre trois espèces d'arbres. En effet, nous pouvons distinguer trois colonnes qui contiennent plus beaucoup de cellules vertes. Ces colonnes portent les labels 9, 43 et 55 qui correspondent respectivement aux espèces *Betula pendula*, *Platanus* x *hispanica* et *Quercus ilex*. Pour ces espèces, nous avons beaucoup d'images d'écorce d'arbres d'âges différents. Si nous prenons l'exemple de *Betula pendula*, l'écorce est blanche et lisse quand l'arbre est jeune, ou avec des lentilles linéaires, et elle devient de moins en moins blanche avec l'âge, pour devenir brune et sillonnée comme beaucoup d'autres arbres. Ainsi, une classe peut contenir des écorces appartenant à la même espèce mais de couleurs différentes et de textures

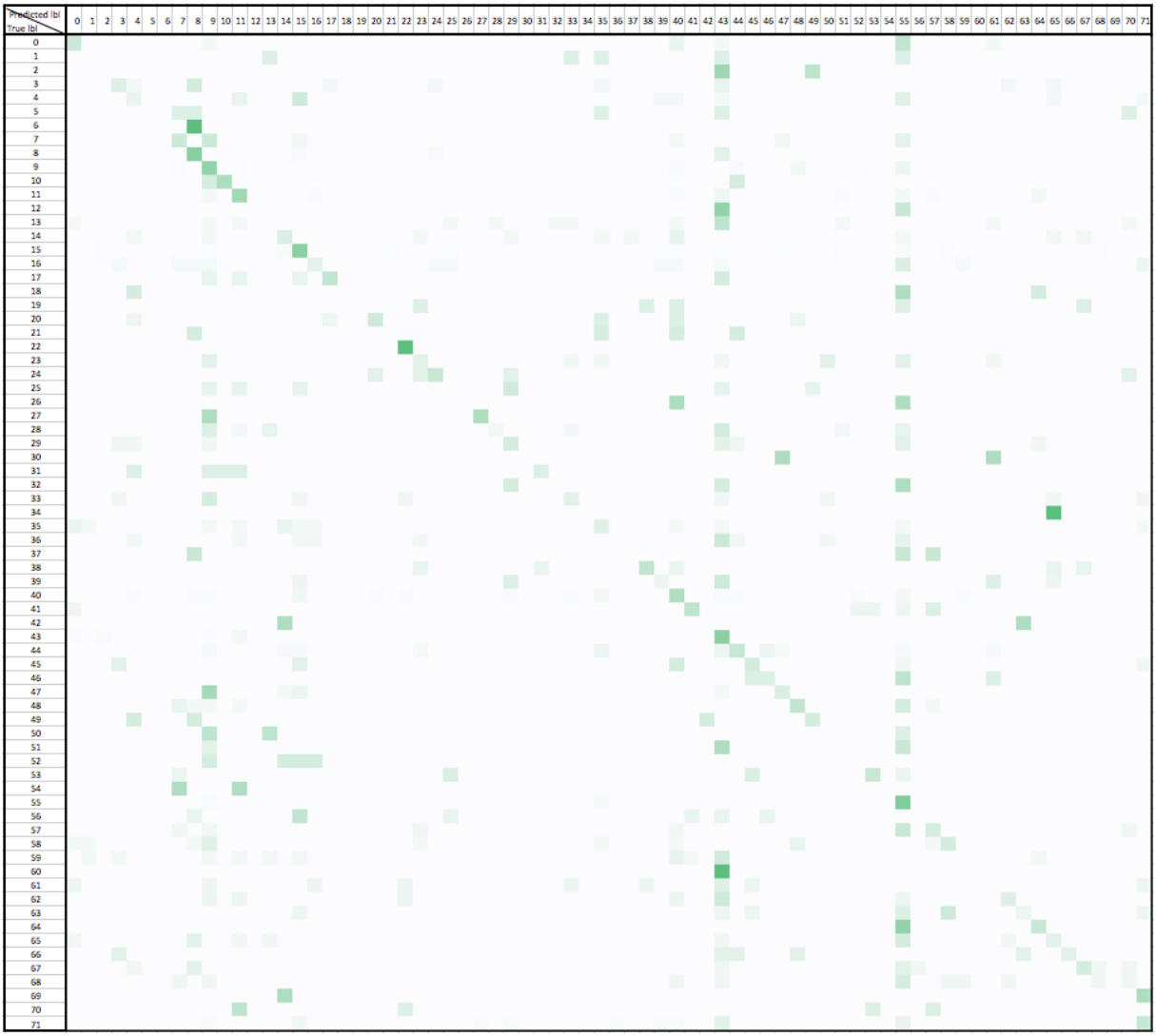


Figure 6.6 – Matrice de confusion issue de la classification des écorces grâce au classifieur SVM un contre tous.

différentes. Cela n'aide pas le classifieur et au contraire, cela ajoute de la confusion.

Concernant les feuilles, ce sont les labels 32, 33 et 55 qui apportent de la confusion. Les espèces correspondantes sont *Ilex aquifolium*, *Laurus nobilis* et *Quercus ilex*. Au sein de ces espèces, la variabilité intra classe est assez élevée, en particulier pour la marge des feuilles de *Ilex aquifolium* (Figure 6.7).

Une autre raison de la présence d'espèces sur-prédites est que la base de données d'apprentissage est déséquilibrée. En effet, nous pouvons avoir plus de 50 échantillons d'écorce pour une espèce (par exemple le label 9, *Betula pendula*, ou le label 45, *Platanus* x *hispanica*) alors que la majorité des espèces n'ont qu'une dizaine d'images d'entraînement. Comme le classifieur a vu plus d'exemples de certaines espèces, il a tendance à prédire ces espèces lorsqu'il n'y a pas de preuve solide que l'échantillon de test appartient à une autre espèce. Le classifieur sur-estime donc ces espèces.

FIGURE 6.7 – Variabilité de la marge des feuilles du houx commun (*Ilex aquifolium* [Cer13b]).

6.4.2 Matrices de confusion pour la concaténation et de la combinaison *a posteriori* des vecteurs de feuilles et des écorces

En ce qui concerne la matrice de confusion pour la concaténation des deux vecteurs sur (Figure 6.8), nous pouvons voir que la confusion entre les espèces diminue, car la diagonale est plus contrastée, et le reste de la matrice ne présente pas de colonnes avec beaucoup de cellules vertes.

En ce qui concerne la combinaison *a posteriori*, nous utilisons les résultats de classification du classifieur de *Folia* pour les caractéristiques des feuilles et le classifieur basé SVM pour les caractéristiques des écorces. En effet, comme présenté dans la partie 6.3, ce sont ces systèmes de classification qui donnent les meilleures performances pour chaque organe. Sur la matrice de confusion pour la combinaison *a posteriori* (Figure 6.9), nous avons quelques confusions par rapport à la feuille seule ou à l'écorce seule, comme pour la combinaison.

Nous pouvons en déduire que cette méthode tend à mieux considérer les éléments caractérisant la feuille quand l'écorce ajoute de la confusion (par exemple le label 9, *Betula pendula*, où l'écorce peut présenter des textures très différentes (voir Figure 6.10) mais que la feuille ne présente pas une grande variabilité), et inversement (par exemple le label 32, *Ilex aquifolium*, qui présente une grande variabilité de feuille (voir Figure 6.7) mais pas d'écorce.)

Cependant, nous pouvons observer que les caractéristiques concaténées sont plus discriminantes que la combinaison *a posteriori*. Cela vient du fait que pour la combinaison *a posteriori* ω_l et ω_f sont fixes et identiques pour toutes les espèces. Pourtant, pour certaines espèces, nous pouvons supposer que ces poids varient selon les espèces. Par exemple, si nous prenons le cas du bouleau verruqueux (*Betula pendula*), dont l'écorce blanche est très caractéristique, le poids affilié à l'écorce pourra être plus important que le poids affilié à l'écorce de l'érable champêtre (*Acer campestre*) dont l'écorce grise et rugueuse est assez commune parmi les espèces d'arbres.

Cependant, pour la méthode de combinaison *a posteriori* des feuilles et des écorces, nous observons une colonne avec beaucoup de confusion dans la matrice. En d'autres mots,

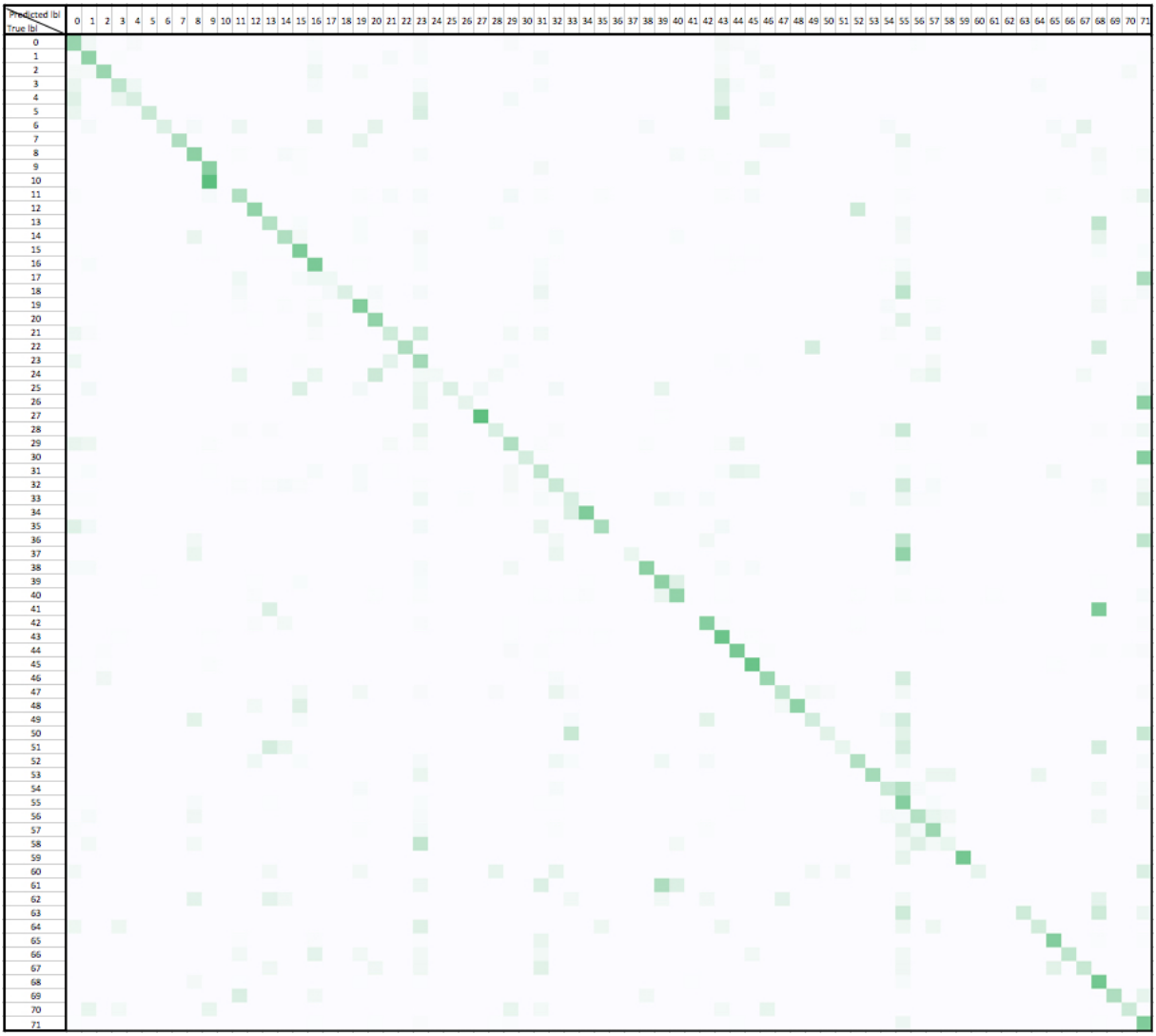


Figure 6.8 – Matrice de confusion issue de la classification avec le classifieur SVM des vecteurs de feuille et d'écorce concaténés.

il y a sur-prédiction d'une classe. Cette classe correspond au label 55 qui est l'espèce *quercus ilex*. Nous pouvons dire que cette méthode de combinaison ne permet pas au système de classification de réduire la sur-prédiction de cette espèce. Ceci vient du fait que *Quercus ilex* est l'espèce pour laquelle il y a de la confusion à la fois dans le système de classification des feuilles et le système de classification des écorces (Figures 6.5 et 6.6). La Figure 6.11 donne une représentation de la grande variabilité des écorces pour l'espèce *Quercus ilex*.

Sur les matrices de confusion pour les méthodes de concaténation et de combinaison *a posteriori* (Figures 6.8 and 6.9), nous pouvons identifier les espèces qui sont confondues avec *Quercus ilex*, notamment *Ilex aquifolium* et *Laurus nobilis*. La Figure 6.12 montre les similitudes qui existent entre les feuilles de ces différentes espèces.

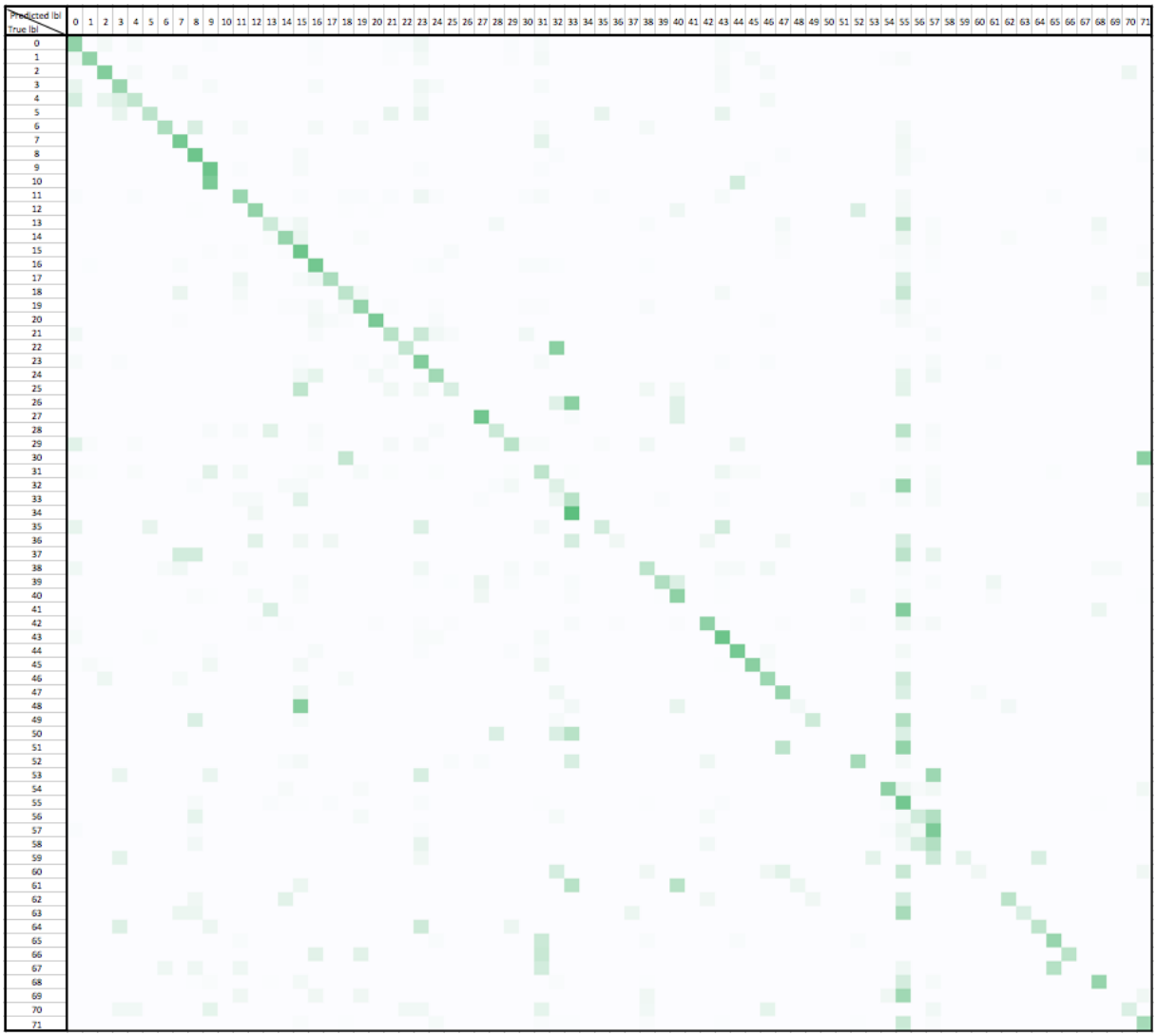


FIGURE 6.9 – Matrice de confusion issue de la combinaison *a posteriori* des vecteurs de feuille et d'écorce.

FIGURE 6.10 – Exemples d'écorces de *Betula pendulas* commençant à un jeune âge et devenant de plus en plus âgées.

6.4.3 Prise en compte des matrices de confusion

L'idée est maintenant de prendre automatiquement en compte ces informations selon lesquelles certaines espèces ajoutent de la confusion afin d'améliorer le système de classification et d'éviter la surestimation pour les espèces mentionnées ci-dessus.

Figure 6.11 – Exemples d'images d'écorces de *Quercus Ilex* présentant une grande variété de textures.

(a) (b) (c) (d)

Figure 6.12 – Confusion entre les feuilles de *Quercus ilex* a) et c) et les feuilles de *Laurus nobilis* b) et *Ilex aquifolium* d).

Premièrement, afin de ne pas biaiser le processus, les matrices de confusion que nous prenons en compte sont calculées sur la base de données d'apprentissage. Comme indiqué précédemment, les matrices de confusion sont normalisées sur les lignes de manière à être invariantes au nombre d'images d'apprentissage de chaque espèce. Toutefois pour être réintroduites dans le processus de classification, celles-ci sont aussi normalisées sur les colonnes dans le but d'être invariantes au nombre de fois qu'une espèce est prédite, donc pour réduire l'impact des classes sur-prédites (Algorithme 2). Par la suite, nous appelons ces matrices "matrices de raffinement".

Algorithm 2 Algorithme de calcul de la matrice de raffinement.

int nbTestSample ← *nombre d'images de test*
for (int a = 0 ; a < nbTestSample ; a++) **do**
  if j == i **then**
    $M_{i,j}++$
  end if
end for
$M_{i,j} \leftarrow M_{i,j} / \sum nbTestSample \in i$ *//normalisation sur les lignes*
$M_{i,j} \leftarrow M_{i,j} / \sum nbTestSample \in j$ *//normalisation sur les colonnes*

Après le processus de classification sur chacun des organes, nous obtenons la prédiction d'appartenance pour chaque classe, qui correspond à une valeur entre 0 et 1. La classe prédite par le système de classification est celle qui a la valeur la plus élevée. Nous utilisons la matrice de raffinement pour pondérer la valeur prédite pour chaque classe et donc pour

réorganiser les classes prédites. Pour ce faire, nous calculons le produit matriciel entre la matrice raffinée M_r et le vecteur v contenant les valeurs d'appartenance à l'étiquette, afin d'obtenir un vecteur réarrangé v_r (équation 6.3).

$$\mathbf{v}_r = \mathbf{v} \cdot M_r \tag{6.3}$$

Raffinement pour les écorces

Dans cette partie nous présentons les résultats de classification pour les vecteurs d'écorces complets **B** : $\boldsymbol{Gab}$, $\boldsymbol{H}_{180}$, $\boldsymbol{W}_H$, $\boldsymbol{W}_V$ et les vecteurs d'écorces réduits **B'** : $\boldsymbol{Gab}$,$\boldsymbol{H}_{30}$, $\boldsymbol{W}_H$, $\boldsymbol{W}_V$. Les résultats sont présentés sur la Figure 6.13.

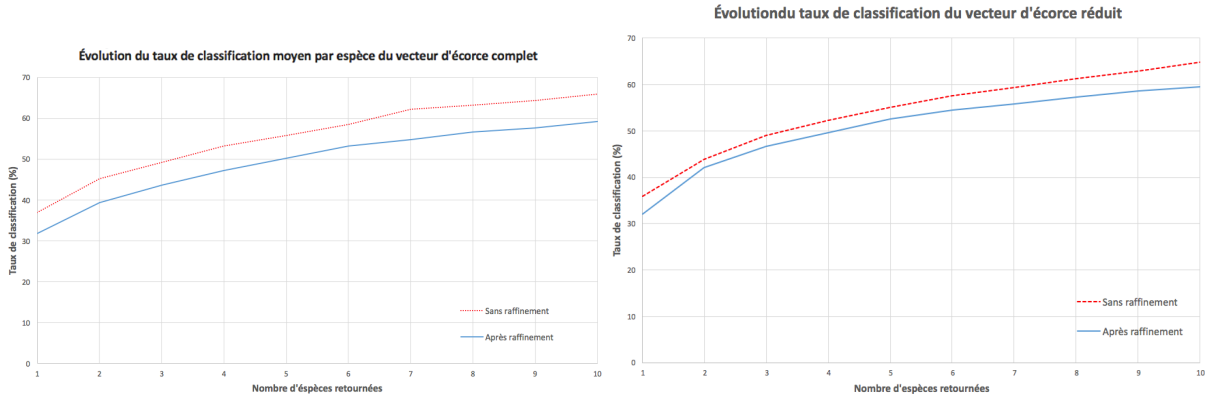


FIGURE 6.13 – Résultats de classification globaux sur les écorces avant et après avoir pris en compte la matrice.

Premièrement, nous pouvons constater que les taux de classification globaux, que ce soit pour le vecteur d'écorce complet ou réduit, sont inférieurs après la prise en compte des matrices de confusion (courbes bleues) par rapport à la classification initiale (courbes rouges). Ceci vient du fait que nous nous mettons à moins bien reconnaître les écorces qui ajoutent beaucoup de confusion, alors que ces espèces ont beaucoup de données de test. Par conséquent, nous nous basons sur le taux de reconnaissance par espèce pour évaluer les performances de cette méthode. L'évolution du taux de classification moyen par espèce est présentée sur la Figure 6.14.

Ainsi, nous pouvons remarquer qu'après la prise en compte de la matrice de confusion (courbe bleue) les taux de reconnaissance par espèce sont supérieurs aux taux obtenus avec la classification initiale. Concrètement, les espèces qui ajoutent de la confusion (classes 9, 43 et 55) sont moins bien identifiées mais toutes les autres espèces voient leurs taux d'identification augmenter. Pour plus de précision, la Figure 6.1 présente les taux de classification du vecteur d'écorce complet pour chaque espèce avant et après prise en compte de la matrice de raffinement.

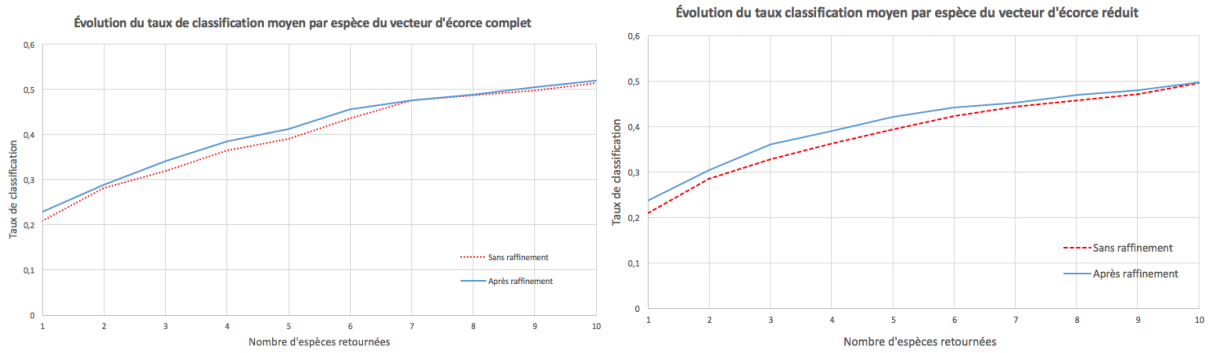


Figure 6.14 – Résultats de classification moyens par espèce sur les écorces avant et après avoir pris en compte la matrice de confusion.

Table 6.1 – Évolution des taux de classification pour chaque espèce avant et après raffinement pour le vecteur d'écorce complet.

Species label	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Identification rate before refinement (%)	31	0	0	20	12	0	0	31	74	67	50	58	0	0	21	72	16	38	0	0	30	0	100	17
Identification rate after refinement (%)	31	0	40	0	18	0	0	38	65	49	50	65	0	28	29	63	16	25	0	0	40	0	100	42
Species label	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
Identification rate before refinement (%)	33	0	0	50	7	25	0	20	0	17	0	19	0	0	38	10	49	40	0	72	35	25	20	22
Identification rate after refinement (%)	33	14	0	50	33	33	0	20	0	25	0	25	0	0	38	10	49	30	0	0	45	25	20	33
Species label	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71
Identification rate before refinement (%)	38	25	0	0	0	33	0	78	0	25	25	0	0	0	20	10	33	17	17	25	8	0	0	35
Identification rate after refinement (%)	50	25	20	33	0	33	0	60	0	33	31	0	0	10	10	20	33	8	0	42	25	0	0	35

Raffinement pour les feuilles

Dans cette partie, nous traitons de la prise en compte de la matrice de confusion pour la classification du vecteur de feuille avec le classifeur de l'application *Folia* qui donne les meilleurs taux sur ces vecteurs. Les résultats de classification avant et après raffinement grâce à la matrice de confusion calculée sur les données d'apprentissage sont présentés sur la Figure 6.15.

Figure 6.15 – Résultats de classification globaux et moyens par espèce sur les feuilles avant et après avoir pris en compte la matrice de confusion.

De même que pour les écorces, nous pouvons observer que les résultats de classification globaux sont inférieurs après raffinement du fait que les espèces les plus présentes dans les bases de données d'apprentissage et de test sont moins reconnues. En revanche, les taux de classification moyens par espèce sont supérieurs de 2% pour le Top 1 et jusqu'à 5% pour le Top 10 après la prise en compte de la matrice de confusion.

Raffinement pour les écorces et feuilles concaténées

Utilisant la même méthode de raffinement, la Figure 6.16 expose les résultats de classification avant et après raffinement pour le vecteur composé du vecteur caractéristique de la feuille et le vecteur caractéristique de l'écorce concaténé pour une liste allant jusqu'à dix espèces proposées. Ici aussi les taux de classification généraux après raffinement (courbe bleue) sont inférieurs aux taux initiaux (courbe rouge). Cependant, après la prise en compte de la matrice de raffinement, les taux d'identification par espèce ne sont pas supérieurs aux taux initiaux (Figure 6.17). L'utilisation de la matrice de confusion n'a pas amélioré la classification des espèces. Cette remarque est valable pour les vecteurs d'écorces complet et réduit.

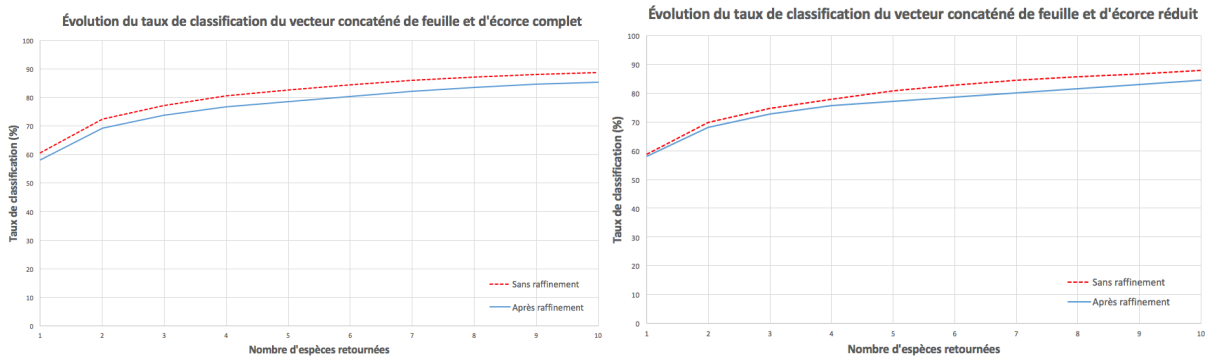


FIGURE 6.16 – Résultats de classification globaux sur les vecteurs de feuilles et d'écorces concaténés avant et après avoir pris en compte la matrice.

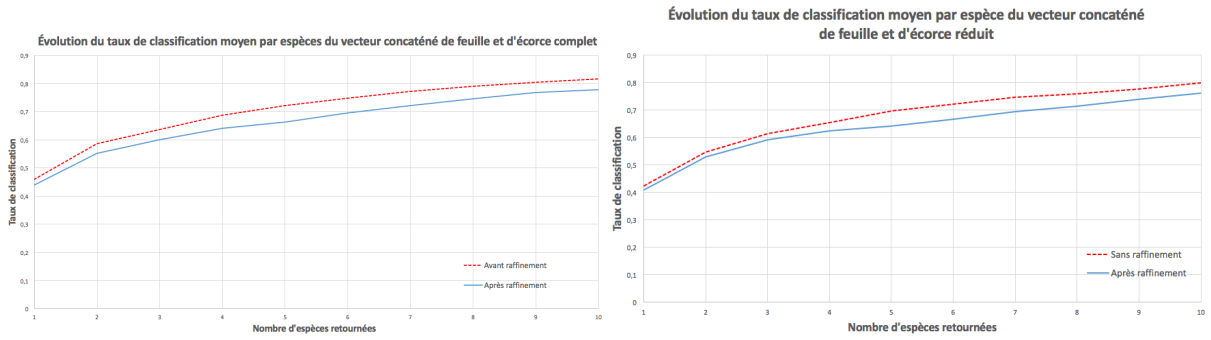


FIGURE 6.17 – Résultats de classification moyens par espèce sur les vecteurs de feuilles et d'écorces concaténés avant et après avoir pris en compte la matrice de confusion.

Raffinement pour les écorces et feuilles combinées *a posteriori*

Concernant la combinaison *a posteriori* le raffinement est effectué avant la combinaison des résultats. Concrètement, un raffinement est fait sur les résultats de classification des feuilles avec la matrice de raffinement des feuilles, et un raffinement est réalisé sur les résultats de classification des écorces avec la matrice de raffinement des écorces. Puis ces résultats sont combinés selon l'Équation 6.2.

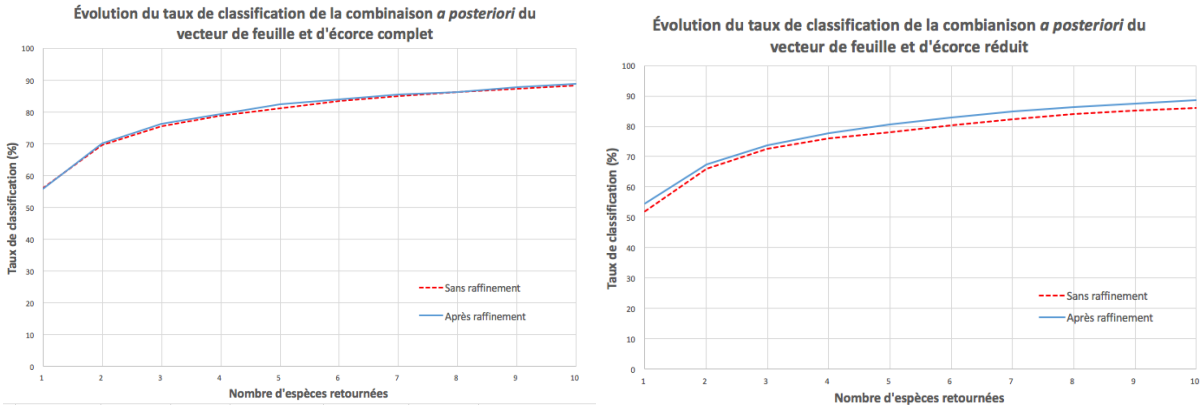


Figure 6.18 – Résultats de classification globaux sur les vecteurs de feuilles et d'écorces combinés *a posteriori* avant et après avoir pris en compte la matrice.

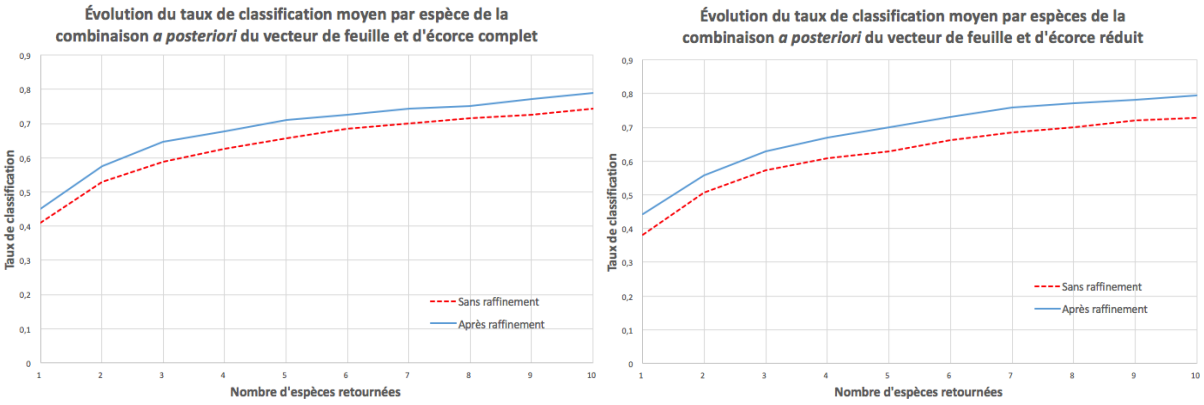


Figure 6.19 – Résultats de classification moyens par espèce sur les vecteurs de feuilles et d'écorces combinés *a posteriori* avant et après avoir pris en compte la matrice de confusion.

Dans cette configuration, nous pouvons voir que raffiner les résultats de classification des feuilles et des écorces donne des taux de classification globaux équivalents lorsque l'histogramme de couleurs est composé de 180 bins et légèrement supérieurs quand l'histogramme est réduit à 30 bins. En revanche, les taux de classification moyens par espèce sont bien supérieurs lorsque les matrices de confusion sont prises en compte.

Dans le tableau suivant, nous proposons de résumer les différents taux de classification globaux et par espèce, selon les différents vecteurs pris en compte et leur combinaison, avec ou sans raffinement.

TABLE 6.2 – Résumé des taux de classification pour les différentes méthodes de combinaison des vecteurs de feuilles et d'écorces.

	Avant raffinement						Après raffinement					
	Taux global (%)			Taux moyen par espèce			Taux global (%)			Taux moyen par espèce		
	Top1	Top5	Top10	Top1	Top5	Top10	Top1	Top5	Top10	Top1	Top5	Top10
Feuille seule	51,1058	80,5687	87,4013	0,357154	0,625704	0,714548	45,3791	77,1327	77,1327	0,377551	0,656636	0,764522
Ecorce seule	36,918	55,765	65,8537	0,209291	0,209291	0,5148	31,8182	50,1109	59,2018	0,229307	0,413192	0,520351
Ecorce réduite seule	35,9202	55,0998	64,8559	0,209046	0,393459	0,495972	32,0399	52,5499	59,5344	0,237271	0,421583	0,496956
Concaténation Feuille + Écorce	**60,6266**	**82,6727**	**88,7505**	**0,460081**	**0,721371**	**0,816966**	**58,0974**	78,5579	85,3907	0,438921	0,662983	0,77771
Concaténation Feuille + Écorce réduite	58,837	80,8722	87,9112	0,424712	0,696049	0,798698	57,9954	77,238	84,5448	0,409302	0,641541	0,761423
Combinaison Feuille + Écorce	55,9296	81,14	88,1408	0,409153	0,657755	0,742877	55,7766	**82,3642**	**88,7529**	**0,449929**	**0,709218**	0,789006
Combinaison Feuille + Écorce réduite	52,0658	78,0796	86,075	0,382073	0,629971	0,729534	54,5524	80,6809	88,4851	0,441606	0,699547	**0,79568**

Pour résumer, la prise en compte de la matrice de confusion sur les données d'apprentissage permet de considérer la confusion qui existe entre les espèces et ainsi aide à augmenter les taux de classification par espèce. Ceci permet d'avoir un système de classification moins dépendant de la densité de population des classes de la base de données, où certaines espèces sont sur-représentées avec plus d'une centaine d'images alors que d'autres en ont à peine une dizaine. Cependant, nous pouvons noter que les meilleurs taux de classification globaux ainsi que par espèce sont obtenus quand les vecteurs de feuilles et d'écorces complets sont simplement concaténés et classifiés avec le classifieur SVM un contre tous. Toutefois, cette méthode de combinaison des données de feuilles et d'écorces nécessite d'embarquer sur le téléphone le classifieur pour les feuilles, le classifieur pour les écorces, et le classifieur pour la concaténation des feuilles et des écorces. Dans le cadre du classifieur SVM un contre tous, il y a un classifieur par espèce stocké sur 250 ko en moyenne, donc pour l'ensemble des espèces il faut multiplier par le nombre de classes (72 dans notre cas). Or les matrices de confusions raffinées ont une taille de 25 ko, ce qui est bien plus faible. Le temps de calcul pour obtenir les résultats de classification raffinés est de moins de 1,9 ns en moyenne, contre 125 ms pour la classification SVM un contre tous sur un MacBook Air avec processeur Intel Core i7 2,2 GHz et 8 Go 1600 MHz de mémoire DDR3. Ainsi, pour des taux légèrement inférieurs, la combinaison *a posteriori* des résultats raffinés permet de réduire le temps de calcul et la taille mémoire du système de classification.

6.5 Ajout des données des fruits et des fleurs

Dans les parties précédentes, nous avons testé des stratégies de fusion entre les données de feuilles et les données d'écorces, car ces deux organes sont les plus présents sur l'arbre tout au long de l'année et sont facilement observables. Cependant, il serait très intéressant de considérer les informations apportées par les fruits ou les fleurs quand ceux-ci sont présents sur l'arbre. Nous allons donc appliquer les méthodes de raffinement et de combinaison *a posteriori* sur ces deux organes.

6.5.1 Raffinement de la prédiction de l'espèce

Pour rappel, la classification des fruits et des fleurs ne se fait pas directement sur l'espèce. Le classifieur MLP (voir Chapitre 5) prédit une classe "*fruitEspeceX*" ou "*fleurEspeceX*". Ici la fusion s'effectue sur les espèces, donc nous utilisons la prédiction sur l'espèce. La matrice de raffinement est calculée sur les données d'apprentissage des fruits et des fleurs, puis est utilisée pour réarranger la prédiction des espèces. La Figure 6.20 présente les résultats de classification avant et après prise en compte de la matrice de raffinement.

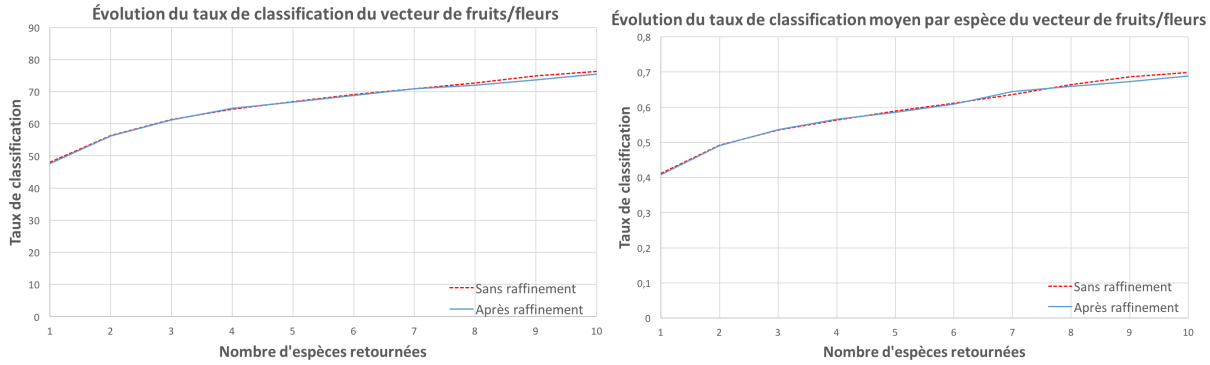


Figure 6.20 – Résultats de classification globaux et moyens par espèce sur les fruits-fleurs avant et après avoir pris en compte la matrice de confusion.

Nous pouvons observer que pour les fruits et les fleurs, la prise en compte de la matrice de raffinement ne permet pas d'augmenter les taux de classification comme cela est le cas pour les feuilles ou les écorces. La raison vient très probablement de l'augmentation de données qui a été réalisée pour l'entraînement des fruits et des fleurs. Cette augmentation de données a permis d'avoir des classes équilibrées, ce qui évite au classifieur de sur-prédire certaines espèces. Ainsi, nous pouvons nous passer de raffinement quand les populations des classes lors de l'apprentissage sont équilibrées.

6.5.2 Combinaison *a posteriori* des fruits-fleurs avec les feuilles et les écorces

Nous nous intéressons à présent à la combinaison *a posteriori* des fruits-fleurs avec les feuilles et les écorces. La concaténation simple des vecteurs n'est pas étudiée car la présence des fruits ou des fleurs est très réduite, ce qui ne justifie pas l'embarquement de trois nouveaux classifieurs pour générer le cas où nous souhaiterions fusionner les feuilles avec les fruits-fleurs, les écorces avec les fruits-fleurs, et les feuilles avec les écorces puis avec les fruits-fleurs. Comme indiqué précédemment, la combinaison *a posteriori* permet d'alléger le système de classification dans l'application.

La Figure 6.21 présente les taux de classification obtenus en combinant *a posteriori* le vecteur de feuille et le vecteur de fruits-fleurs.

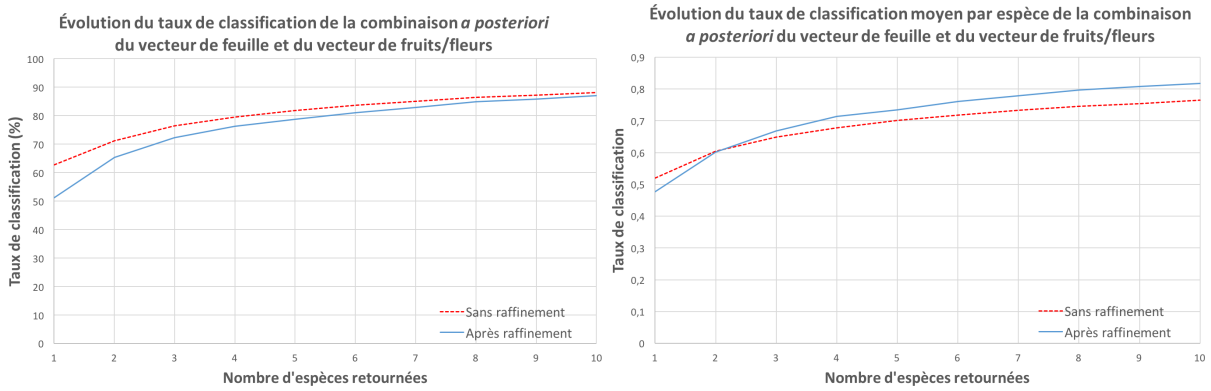


FIGURE 6.21 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion.

Ces résultats sont semblables à ceux que nous avons obtenus précédemment, avec une diminution du taux de classification global mais une augmentation du taux de bonne classification par espèce. Seul le taux de classification moyen par espèce pour le top 1 est meilleur quand la matrice de confusion n'est pas prise en compte.

Deuxièmement, nous effectuons la combinaison *a posteriori* entre les résultats de classification du vecteur d'écorce complet et les résultats de classification des vecteurs de fruits et de fleurs. La Figure 6.22 présente l'évolution des taux de bonne classification de la combinaison de ces organes avant et après raffinement.

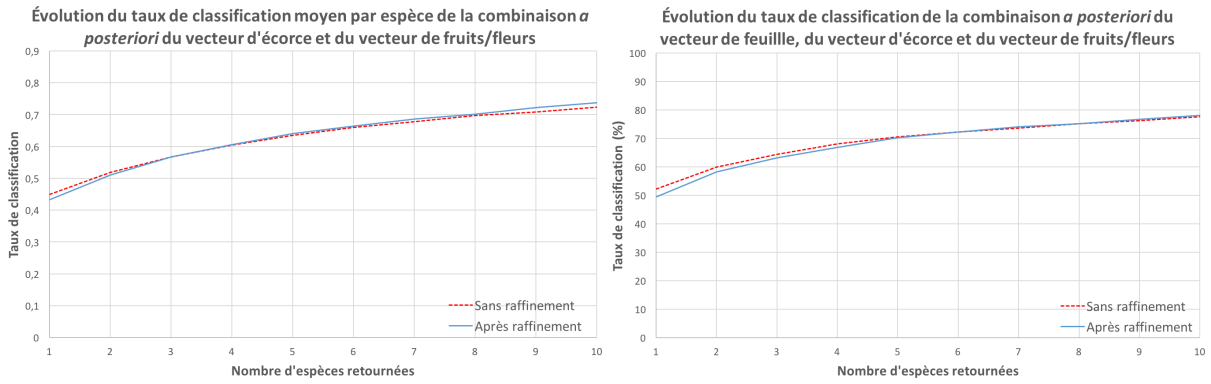


FIGURE 6.22 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des écorces et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion.

Dans le cas de la combinaison *a posteriori* des écorces et des fruits-fleurs, le raffinement n'apporte pas une grosse amélioration des résultats

Troisièmement, nous proposons de faire la combinaison *a posteriori* entre les résultats de classification des feuilles, des écorces, et des fruits-fleurs (Figure 6.23).

Dans cette configuration, la prise en compte de la matrice de confusion améliore les résultats globaux et par espèce. Les matrices de raffinement des différents organes semblent se compléter de manière à faire plus ou moins confiance à la prédiction d'appartenance à

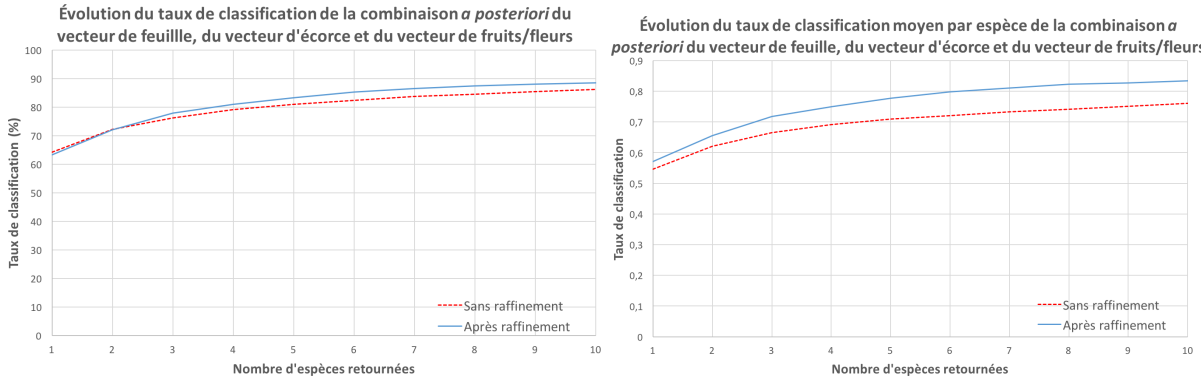


Figure 6.23 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles, des écorces et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion.

l'espèce en fonction de l'organe considéré.

Le Tableau 6.3 permet de résumer les différents taux de bonne classification selon les différentes combinaisons.

Table 6.3 – Résumé des taux de classification pour les différentes méthodes de combinaison des vecteurs de feuilles, d'écorces et de fruits-fleurs.

	Avant raffinement						Après raffinement					
	Taux global			Taux moyen par espèce			Taux global			Taux moyen par espèce		
	Top1	Top5	Top10	Top1	Top5	Top10	Top1	Top5	Top10	Top1	Top5	Top10
Feuille seule	51,1058	80,5687	87,4013	0,35715	0,6257	0,71455	45,3791	77,1327	77,1327	0,37755	0,65664	0,76452
Ecorce seule	36,918	55,765	65,8537	0,20929	0,20929	0,5148	31,8182	50,1109	59,2018	0,22931	0,41319	0,52035
Ecorce réduite seule	35,9202	55,0998	64,8559	0,20905	0,39346	0,49597	32,0399	52,5499	59,5344	0,23727	0,42158	0,49696
Fruits-Fleurs seuls	48,0241	66,9244	76,3746	0,41126	0,58877	0,6986	47,6804	66,7526	75,5155	0,40807	0,58535	0,68832
Concaténation Feuille + Écorce	60,6266	**82,6727**	**88,7505**	0,46008	**0,72137**	**0,81697**	58,0974	78,5579	85,3907	0,43892	0,66298	0,77771
Concaténation Feuille + Écorce réduite	58,837	80,8722	87,9112	0,42471	0,69605	0,7987	57,9954	77,238	84,5448	0,4093	0,64154	0,76142
Combinaison Feuille + Écorce	55,9296	81,14	88,1408	0,40915	0,65776	0,74288	55,7766	82,3642	88,7529	0,44993	0,70922	0,78901
Combinaison Feuille + Écorce réduite	52,0658	78,0796	86,075	0,38207	0,62997	0,72953	54,5524	80,6809	88,4851	0,44161	0,69955	0,79568
Combinaison Feuille + Fruits-Fleurs	62,8261	81,8841	88,0797	0,52031	0,70145	0,76477	51,2681	78,6957	87,1014	0,47655	0,73515	0,81778
Combinaison Écorce + Fruits-Fleurs	52,2398	70,6192	77,6021	0,44942	0,63524	0,72372	49,5389	70,2899	78,1291	0,433	0,63999	0,73732
Combinaison Feuille + Écorce + Fruits-Fleurs	**64,2299**	80,9439	86,2669	**0,54538**	0,70837	0,75981	**63,2363**	**83,286**	**88,5025**	**0,5708**	**0,77688**	**0,83355**

Dans un premier temps nous pouvons remarquer que combiner les fruits-fleurs avec la feuille ou l'écorce permet bien d'augmenter les taux de bonne classification, que ce soit avant ou après raffinement. Ainsi, les informations extraites sur les fruits ou les fleurs n'apportent pas de bruit. Dans aucun cas la combinaison de deux organes ou plus ne fait baisser les chances de bonne identification de l'espèce d'arbre. C'était précédemment la concaténation des vecteurs de feuille et d'écorce qui donnaient les meilleurs résultats de classification. Si nous considérons le fruit ou la fleur, c'est à présent la combinaison *a posteriori* des résultats de classification des feuilles, des écorces et des fruits-fleurs qui permet d'avoir le plus de chance d'identifier la bonne espèce. Si en plus nous utilisons la confusion entre les espèces, nous avons 57,08% de chance d'identifier la bonne espèce et jusqu'à 83,36% de chance d'avoir la bonne espèce présente dans une liste de dix espèces proposées. Nous pouvons dire que ces résultats sont respectables au vu de la faible dimension des vecteurs utilisés et de la complexité des systèmes de classification.

6.5.3 Prise en compte du genre et non de l'espèce

Précédemment la prise en compte du fruit ou de la fleur s'est faite au niveau de l'espèce. Nous avons cependant vu que les fruits et les fleurs d'espèces appartenant à un même genre étaient semblables (cas du gland pour le genre *Quercus* par exemple). C'est pour cette raison que nous proposons dans cette partie de privilégier la prédiction sur le genre plutôt que sur l'espèce pour les informations issues de la classification des fruits et des fleurs.

Nous allons de nouveau utiliser notre méthode de combinaison *a posteriori* (Équation 6.2). Cependant, pour le vecteur de fruits-fleurs, pour chaque espèce ce n'est pas la prédiction d'appartenance à l'espèce qui est associée, mais la prédiction d'appartenance au genre. Par exemple, toutes les espèces du genre *Acer* (*Acer campestre*, *Acer monspessulanum*, *Acer opalus*, *Acer platanoides*, *Acer pseudoplatanus*, et *Acer saccharinum*) auront la même prédiction d'appartenance. La prédiction d'appartenance retenue est la plus élevée. Avec notre exemple, si la prédiction d'appartenance à l'espèce *Acer campestre* est de 0,3 et la prédiction d'appartenance à l'espèce *Acer saccharinum* est de 0,1, la prédiction de l'espèce *Acer saccharinum* devient 0,3.

La Figure 6.24 montre l'évolution du taux de classification global et par espèce des feuilles avant et après la combinaison *a posteriori* des prédictions sur les genres *via* les fleurs et les fruits. La Figure 6.25 montre l'évolution du taux de classification global et par espèce des écorces avant et après la combinaison *a posteriori* des prédictions sur les genres *via* les fleurs et les fruits. Et la Figure 6.26 montre l'évolution du taux de classification global et par espèce de la combinaison *a posteriori* des feuilles et des écorces avant et après la combinaison *a posteriori* des prédictions sur les genres *via* les fleurs et les fruits.

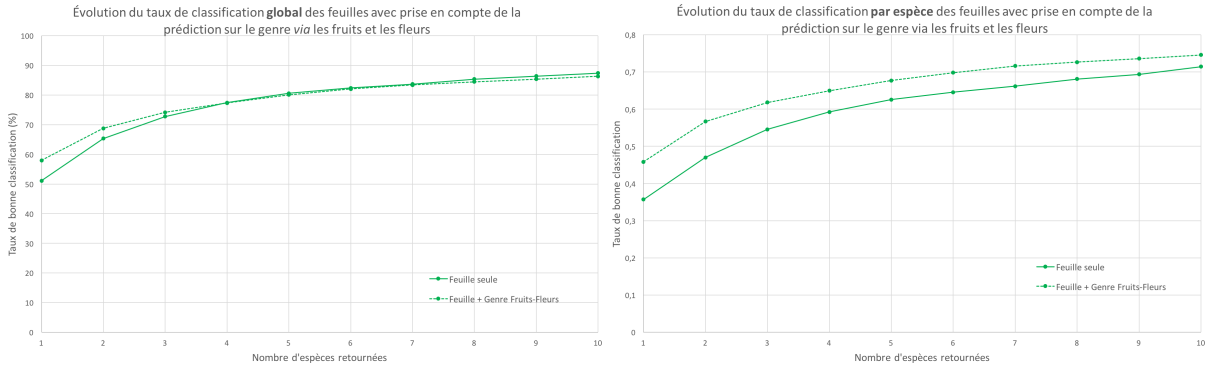


FIGURE 6.24 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres.

Pour les feuilles et la combinaison *a posteriori* des feuilles et des fruits, l'ajout de l'information du genre obtenue grâce aux fleurs et aux fruits permet d'augmenter le taux de bonne classification global pour le Top 1. Elle augmente également le taux de bonne classification par espèce pour une liste allant d'une à dix espèces pour les feuilles et pour

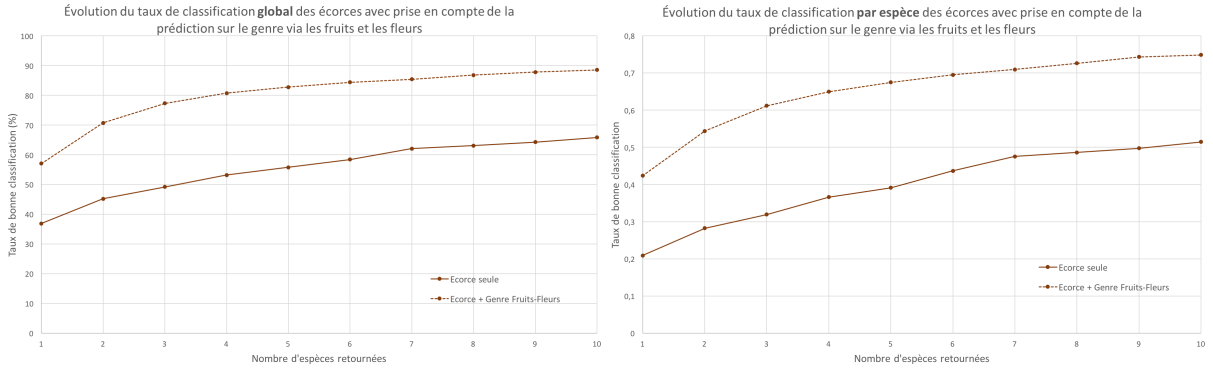


Figure 6.25 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des écorces et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres.

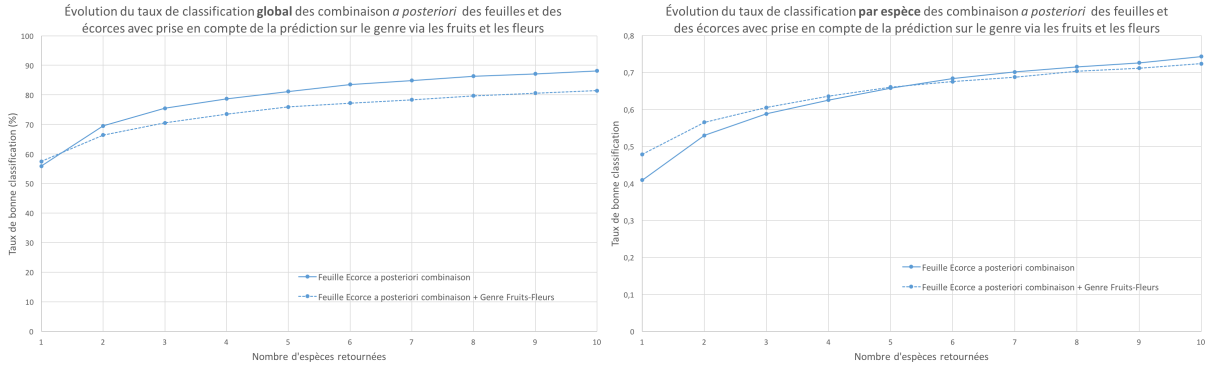


Figure 6.26 – Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles, des écorces et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres.

une liste allant d'une à 5 espèces pour la combinaison des feuilles et des écorces. C'est cependant pour la classification des écorces que l'apport de la prédiction du genre a une grande influence. En effet, les taux de classification sont en moyenne augmentés de 25%.

Le Tableau 6.4 permet de comparer les taux de classification avec la combinaison *a posteriori* des fruits et des fleurs avec la prédiction sur l'espèce ou sur le genre.

Table 6.4 – Résumé des taux de classification pour les combinaisons *a posteriori* des vecteurs de feuilles, d'écorces et de fruits-fleurs en prenant en compte la prédiction sur l'espèce ou sur le genre.

	Taux global			Taux moyen par espèce		
	Top1	Top5	Top10	Top1	Top5	Top10
Combinaison Feuille + Fruits-Fleurs **Espèce**	**62,8261**	**81,8841**	**88,0797**	**0,520311**	**0,701454**	**0,764766**
Combinaison Écorce + Fruits-Fleurs **Espèce**	52,2398	70,6192	77,6021	**0,449418**	0,635236	0,723718
Combinaison Feuille + Écorce + Fruits-Fleurs **Espèce**	**64,2299**	**80,9439**	**86,2669**	**0,545375**	**0,70837**	**0,759806**
Combinaison Feuille + Fruits-Fleurs **Genre**	57,971	80,0725	86,3044	0,458572	0,625704	0,714548
Combinaison Écorce + Fruits-Fleurs **Genre**	**57,0773**	**82,785**	**88,5616**	0,423838	**0,674698**	**0,748595**
Combinaison Feuille + Écorce + Fruits-Fleurs **Genre**	57,4521	75,9049	81,4053	0,478648	0,66058	0,723789

D'après le Tableau 6.4, nous pouvons conclure que l'information de genre n'est pas plus

avantageuse que l'information d'appartenance à l'espèce simple. Le classifieur MLP utilisé pour la classification des fruits et des fleurs arrive à faire la distinction entre les fruits et les fleurs d'espèces appartenant à un même genre malgré leur similitude. Il y a tout de même une exception pour la combinaison *a posteriori* de l'écorce avec les fruits-fleurs, où dans ce cas l'information de genre obtenue avec les fruits-fleurs est plus pertinente que l'information d'espèce.

6.6 Conclusion

Nous avons présenté dans ce chapitre différentes stratégies de fusion de données pour combiner les informations issues des différents organes des arbres que sont les feuilles, les écorces, les fruits et les fleurs. L'identification de l'espèce d'un arbre peut être difficile même pour un botaniste à partir uniquement d'une photo de feuille, et encore plus complexe pour une photo d'écorce [WSG11]. Or, nous avons montré que considérer plusieurs organes permet d'augmenter les taux de bonne identification des espèces d'arbres. Les meilleurs taux sont obtenus quand la feuille, l'écorce et le fruit ou la fleur sont disponibles et leur reconnaissance combinée *a posteriori*.

Par ailleurs, nous avons proposé une méthode permettant de prendre en compte la confusion qui existe entre les espèces ou induite par le classifieur grâce aux matrices de confusion. La prise en compte de la matrice de confusion permet de raffiner les résultats et ainsi d'augmenter les taux de reconnaissance par espèce. De cette manière, pour la combinaison *a posteriori* de la feuille, de l'écorce, et du fruit ou de la fleur, le taux de bonne reconnaissance par espèce passe de 54,54% à 57,08% pour le Top 1. Dans le cas de l'application *Folia* qui propose une liste de dix espèces probables, il y a 83,34% de chance d'avoir la bonne espèce proposée quand les trois organes sont disponibles et que la confusion est prise en compte. La prise en compte de la matrice de confusion est une stratégie qui n'alourdit pas excessivement le système de classification en vue d'une embarcation sur smartphone.

Pour finir, nous avons proposé de ne pas considérer l'information de prédiction sur l'espèce grâce aux fruits et aux fleurs mais plutôt la prédiction sur le genre en raison de l'information de clade qu'apportent ces organes. Nous avons montré que l'information du genre à la place de l'espèce n'est bénéfique que pour la combinaison avec l'écorce. Pour la combinaison *a posteriori* avec les feuilles, il vaut mieux privilégier la prédiction sur les espèces.

Chapitre 7

Interprétation sémantique des informations extraites par un réseau convolutionnel

Jusqu'à présent, les méthodes mises en place pour la reconnaissance d'espèces d'arbres utilisent des descripteurs choisis et développés "à la main" afin d'extraire des informations botaniques des différents organes de la plante (écorce, fruit, fleur). Cependant, l'état de l'art fait de plus en plus usage des réseaux de neurones à convolutions (CNN) pour l'identification des espèces végétales [Jol+16b ; Jol+17]. Ces outils sont largement utilisés de par leur grande efficacité à partir du moment où le nombre de données est suffisant. En outre, en plus du système de classification, les CNN permettent d'extraire automatiquement les caractéristiques pertinentes sur les images pour leur classification. Cela supprime l'étape du choix des descripteurs optimaux au concepteur. Néanmoins, ces CNN fonctionnent encore un peu comme des "boîtes noires", sans possibilité de savoir ce qui est extrait sur les images. Dans ce chapitre, nous cherchons à voir si ces caractéristiques extraites par les filtres des CNN portent une information botanique. Le but serait à terme de pouvoir profiter de la puissance de classification des CNN tout en retournant les critères botaniques à l'utilisateur (forme des organes, couleurs, *etc.*).

Les parties 7.3 et 7.4 présentent les travaux menés conjointement avec Frédéric PHAN, stagiaire de Master 1 au LIRIS d'avril à juillet 2018.

7.1 Cadre de travail

Le réseau de neurones à convolution de base est le perceptron multi-couche, ou MLP. Les MLP sont efficaces pour classifier les données mais ceux-ci sont limités quand la taille des données d'entrée est trop importante, du fait que chaque neurone de chaque couche est entièrement connecté à chacun des neurones de la couche suivante. Les réseaux de neurones à convolutions permettent de réduire la taille des données en entrée du MLP. Un CNN est composé de plusieurs couches dont des couches de convolutions, des couches de pooling, des couches de correction et des couches entièrement connectées (MLP). La couche de convolution (CONV) est l'élément central des CNN. Elle est composée d'un ensemble de filtres de convolution. Lors de la phase d'entraînement, le CNN apprend les filtres pertinents qui détectent un type de caractéristique spécifique à une certaine position spatiale dans l'entrée. La couche de pooling (POOL) permet de réduire la dimensionnalité des données, généralement en effectuant un sous échantillonnage. La couche de correction (ReLU) permet d'appliquer une fonction d'activation sur les signaux en entrée. Enfin, la couche entièrement connectée (FC) permet de prendre une décision de classification selon les données transformées par les différentes couches précédentes.

7.1.1 Réseau de neurones à convolutions utilisé

Il existe un certain nombre d'architectures de réseaux de neurones, de plus en plus complexes et performantes qui, chaque année, permettent d'obtenir de meilleurs taux de

classification sur de grandes bases de données d'images telles qu'ImageNet [1]. ImageNet comporte plus de 15 millions d'images réparties dans plus de 20 000 catégories. Le challenge ILSVRC, permettant aux équipes de recherche de comparer leurs algorithmes de classification, utilise un sous-ensemble d'ImageNet avec environ 1000 images dans chacune des 1000 catégories. En tout, il y a environ 1,2 million d'images de d'apprentissage, 50 000 images de validation et 150 000 images de test. Les CNN obtenant actuellement les meilleures performances, tels que VGG, GoogLeNet, ResNet, SqueezeNet, *etc.*, contiennent cependant un grand nombre de couches et une architecture complexe. Par ailleurs, il n'existe actuellement pas, à notre connaissance, de base de données d'organes végétaux permettant d'entraîner correctement ces types de CNN.

Un des premiers CNN à avoir obtenu de bonnes performances sur le jeu de données ImageNet est le réseau AlexNet [KSH12]. Ce réseau atteint des taux d'erreur de 37,5% pour le top 1 et 17,0% pour le top 5, contre 45,7% et 25,7% respectivement lors des éditions précédentes du challenge ILSVRC [KSH12]. Ce réseau à l'avantage de présenter une architecture linéaire, avec un nombre de couches raisonnable afin d'être étudiées et interprétées. En effet, celui-ci ne comporte que 5 couches de convolution (Figure 7.1).

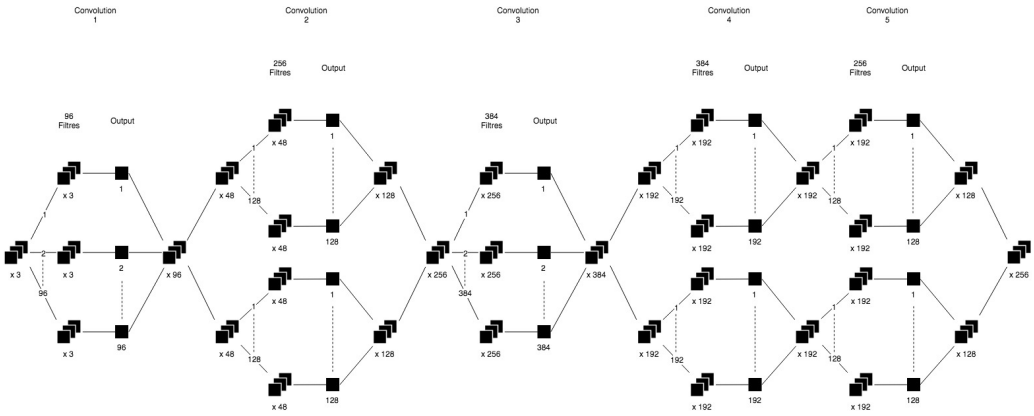


Figure 7.1 – Illustration par Frédéric PHAN des couches de convolutions du CNN AlexNet.

En entrée du réseau, nous fournissons une image couleur, donc composée de trois canaux R, G, B. La première couche de convolution contient 96 filtres. Dans la deuxième couche de convolution, les sorties de la première couche sont séparées en deux groupes afin d'entraîner les données sur deux GPU séparés et donc d'accélérer l'apprentissage. Chaque groupe a une image de 48 canaux en entrée, qui va être filtrée par 128 filtres. Nous pouvons remarquer que les sorties de la deuxième couche de convolution sont regroupées dans une image de 256 canaux, en effet cela permet au réseau d'apprendre les caractéristiques extraites sur les deux GPU. Dans la troisième couche de convolution, le processus de convolution est le même que celui de la première couche de convolution, à la différence du nombre de filtres utilisés qui est de 384. Les résultats sont alors de nouveau séparés en deux

1. https ://www.image-net.org/challenges/LSVRC

groupes pour les couches quatre et cinq afin de faire le même processus de convolution que celui de la deuxième couche de convolution. À la sortie de la cinquième couche de convolution, nous obtenons 256 sorties filtrées.

7.1.2 Données utilisées

Interpréter les résultats de visualisation des différents filtres de convolution peut être complexe pour un utilisateur humain. Afin de faciliter ce travail, nous avons choisi d'utiliser des images botaniques simples. En l'occurrence, nous nous sommes orienté sur des images où l'objet d'intérêt est déjà extrait de son environnement. De plus, le but étant de déterminer si un CNN extrait des informations botaniques des images qui lui sont données, nous avons choisi un organe facilement descriptible à partir d'une simple photo : la feuille. Les aspects morphologiques de la feuille portant une information botanique sont la forme générale du limbe, la forme de l'apex, la forme de la base, et la description de la marge (le bord).

Une partie de la base PlantCLEF est un jeu de données intitulé "Leaf (scan and scan-like)" [Goë+14]. Cette base de données contient des photos de feuilles photographiées sur un fond uni contrastant avec la couleur de la feuille. Les feuilles sont scannées ou photographiées à plat. De ce fait, il peut y avoir des reflets ou des ombres sur les images. Cependant, il n'y a qu'une seule feuille simple par image et celle-ci est centrée, avec l'apex vers le haut. La Figure 7.2 présente des exemples d'images de feuilles issues de ce jeu de données.

FIGURE 7.2 – Exemples d'images de la partie "Leaf (scan and scan-like)" de la base de données PlantCLEF.

Notons toutefois que le réseau de neurones à convolutions ne peut être directement entraîné sur ces images, car les dimensions de celles-ci ne sont pas constantes. En effet, en entrée du réseau AlexNet, les images doivent être de taille 227x227. Un recadrage est donc nécessaire. Les images ne peuvent pas être recadrées pour que la largeur et la hauteur correspondent à la taille requise car cela déformerait la forme générale de la feuille, notamment le *ratio* entre la longueur et la largeur. Les images ne peuvent pas non plus être découpées selon la dimension la plus faible car cela risquerait de couper la base ou l'apex de la feuille. Une des solutions souvent adoptée en deep learning est de rajouter du bruit uniforme sur les bords de l'image afin de conserver les aspects des objets. La Figure 7.3 montre les images obtenues avec cette méthode.

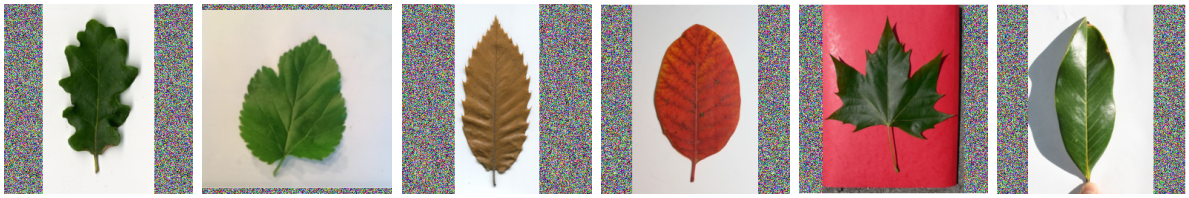


Figure 7.3 – Images de la partie "Leaf (scan and scan-like)" de la base de données PlantCLEF avec du bruit gaussien sur les bords de manière à conserver le *ratio* de la forme des feuilles.

Nous avons sélectionné les images de feuilles des 72 espèces d'arbres qui ont été utilisées au cours de cette thèse. Le réseau de neurones à convolutions AlexNet a été entièrement entraîné sur ces 72 classes (pas de fine-tuning). La base de données est constituée de 7886 images dont 65% est utilisé pour l'entraînement du réseau, 25% pour la validation du modèle, et 10% pour les tests.

7.1.3 Performance du CNN

Le but de ce chapitre n'est pas de créer un CNN pour obtenir les meilleurs résultats de classification. Cependant, afin de s'assurer que le réseau a généré des filtres pertinents et caractéristiques pour reconnaître les espèces d'arbres à partir de photos de feuilles, nous avons mesuré les performances de celui-ci. Si le taux de bonne classification est suffisamment supérieur au hasard alors des caractéristiques appropriées ont été extraites. Ainsi, le taux de bonne classification est de 70,23% pour le Top 1 et de 93,03% pour le Top 5. Les taux de classification sont donc tout à fait corrects, et AlexNet est un bon réseau de neurones à convolution pour l'identification d'espèces d'arbres à partir d'une image de feuille simple photographiée sur fond uni.

7.2 Visualisation des couches

7.2.1 Principes et fonctionnement

Nous aimerions, pour chaque couche de convolution, visualiser les sorties des différents filtres pour interpréter ce que ceux-ci extraient [Sel+17]. La première couche de convolution du réseau AlexNet est constituée de 96 filtres de taille 11x11. Chaque filtre est composé de 3 canaux, car l'image en entrée est en couleur (RGB). Pour chaque filtre, nous effectuons la convolution entre l'image en entrée du réseau et le filtre, et ce pour chaque canal. Pour finir, une moyenne pondérée est réalisée entre les canaux afin de retourner une seule image par filtre. En sortie de la première couche de convolution, nous avons donc autant d'images que de filtres, soit 96. Afin de faciliter la visualisation des images en sortie de couche, il est nécessaire de normaliser les images convoluées. La valeur minimale et la valeur maximale de toutes les images convoluées sur une couche sont

récupérées puis utilisées pour transposer les valeurs des pixels des images entre -1 et 1. Aussi, les images sont représentées avec un mappage de couleur "Jet" (Figure 7.4). Nous utilisons une fonction de base triangulaire tronquée définie comme suit :

$$f_{[0,1]}(x) = \begin{bmatrix} 0 \text{ si } x < 0 \\ x \text{ si } 0 < x < 1 \\ 1 \text{ si } x > 1 \end{bmatrix} \tag{7.1}$$

La palette de couleur "Jet" définit les couleurs RGB selon les équations suivantes :

$$\begin{aligned} R(x) &= f_{[0,1]}(1,5 - |2x - 1|) \\ V(x) &= f_{[0,1]}(1,5 - |2x|) \\ B(x) &= f_{[0,1]}(1,5 - |2x + 1|) \end{aligned} \tag{7.2}$$

Si nous traçons les évolutions des valeurs des trois composantes colorimétriques, nous obtenons les courbes présentes sur la Figure 7.4 et la combinaison de ces trois composantes donne le mappage "Jet".

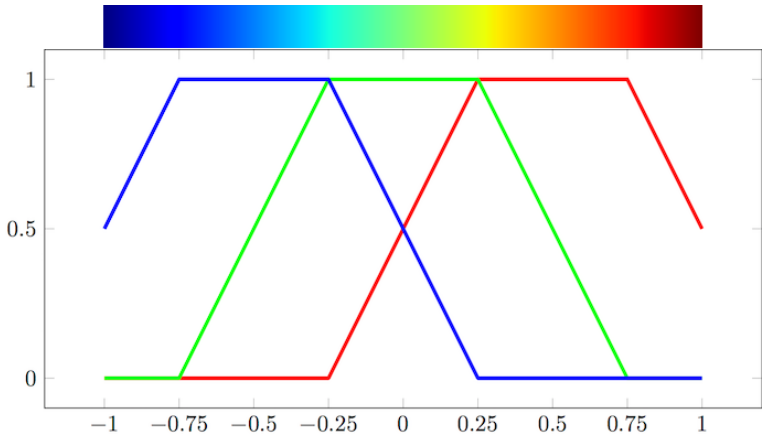


FIGURE 7.4 – Échelle de couleur "Jet".

En entrée de la deuxième couche de convolution, nous avons les 96 images résultantes des convolutions de la première couche. Ces 96 images sont en fait considérées comme une seule image avec 96 canaux. La deuxième couche de convolution du réseau AlexNet comporte 256 filtres. De manière similaire à la première couche de convolution, chacun des 256 filtres possède 96 canaux. Une fois les convolutions effectuées sur tous les canaux, une moyenne pondérée est de nouveau effectuée entre les canaux de manière à se retrouver avec une seule image avec un seul canal par filtre. En sortie de la deuxième couche de convolution, nous avons une image de 256 canaux (car 256 filtres sur la deuxième couche). Ce processus est répété sur les trois dernières couches de manière similaire, avec 384 filtres sur la troisième et quatrième couche de convolution, et 256 filtres sur la cinquième couche.

7.2.2 Visualisation préliminaire

Grâce à cette méthode de visualisation, nous pouvons étudier les parties des images qui sont considérées comme pertinentes par le CNN AlexNet. Dans un premier temps, nous essayons de repérer visuellement quels sont les filtres qui s'activent quand des feuilles possédant une caractéristique botanique identique sont données en entrée du réseau et passent à travers les différentes couches de convolutions. Un filtre est considéré comme "actif" quand le résultat de la convolution entre l'image en entrée et le noyau de convolution contient des valeurs de pixels non proches de zéro. Nous avons fait une sélection de feuilles présentant des caractéristiques botaniques similaires ou dissemblables (Figure 7.5).

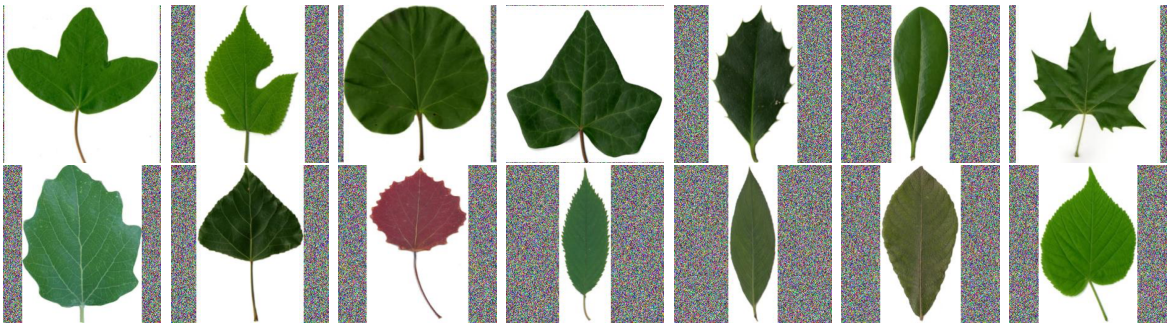


Figure 7.5 – Sélection de feuilles pour visualiser les différentes couches de convolutions.

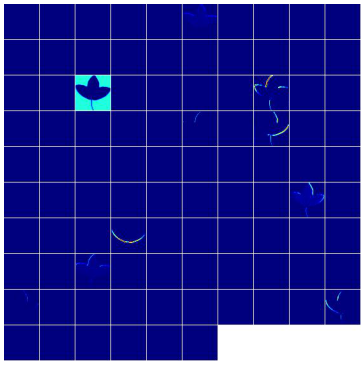
Les images représentent des feuilles qui ont un apex pointu ou plutôt arrondi, des bases plates, aigües ou obtuses, des formes allongées, plutôt larges, ou lobées, et enfin avec des bords lisses ou dentées. Chaque image est donnée au réseau pour classification, et les images convoluées par les filtres des cinq couches de convolution sont sauvegardées. Un exemple de visualisation pour chaque couche est donné (Figure 7.6).

En premier lieu sur les résultats de convolution de la première couche, nous pouvons remarquer que certains filtres s'activent et que d'autre non. La Figure 7.7 présente les filtres actifs et non actifs de la première couche de convolution après le passage de l'image de feuille présentée sur la Figure 7.6a.

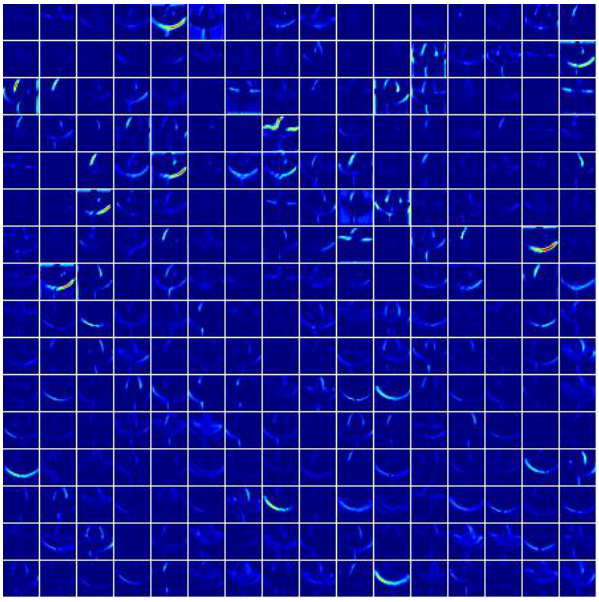
En étudiant les filtres des deux premières couches de convolutions, nous pouvons remarquer que ce sont principalement les informations de bas niveau telles que les contours qui sont extraits selon plusieurs orientations. Nous pouvons aussi remarquer qu'il y a des filtres qui isolent des contours ayant la même orientation, mais avec plus ou moins d'intensité. Certains sont très peu activés. La Figure 7.8 montre pour trois feuilles différentes les résultats de trois filtres identiques de convolutions sur la deuxième couche.

Nous pouvons donc remarquer que pour un filtre particulier de la première ou de la deuxième couche de convolution, celui-ci peut mettre en évidence les contours d'une feuille selon une orientation bien précise. À partir des informations de bas niveau qui sont extraites sur ces deux premières couches, ces données peuvent être combinées par les filtres de convolutions des couches suivantes afin d'extraire des caractéristiques de plus

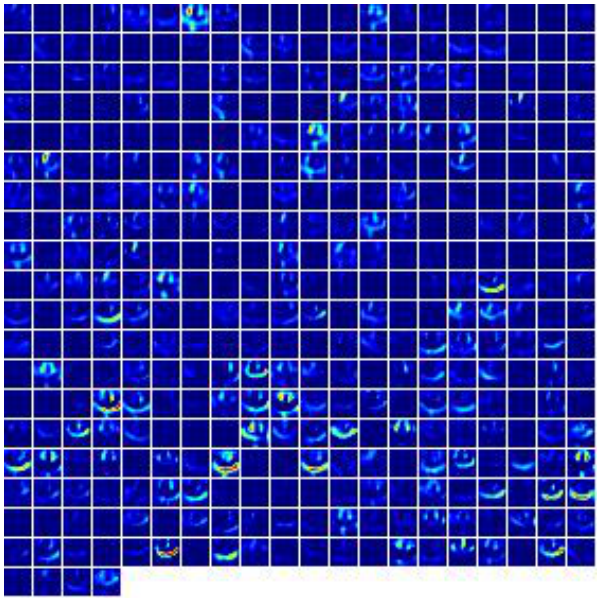
(a) Image mise en entrée du CNN.



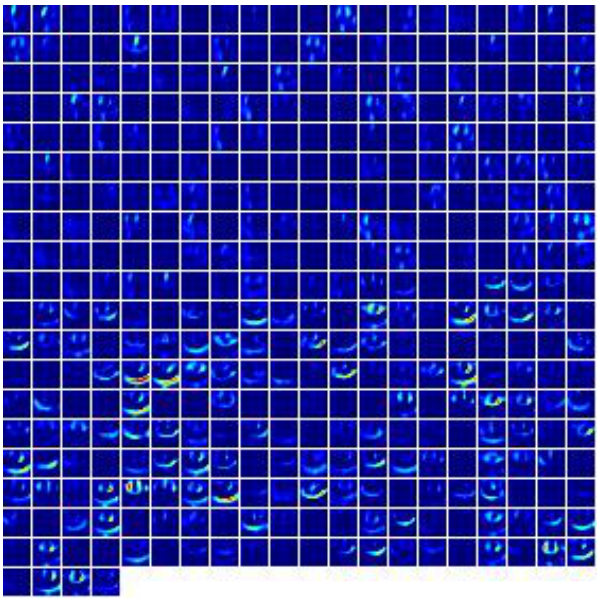
(b) Couche de convolution 1.



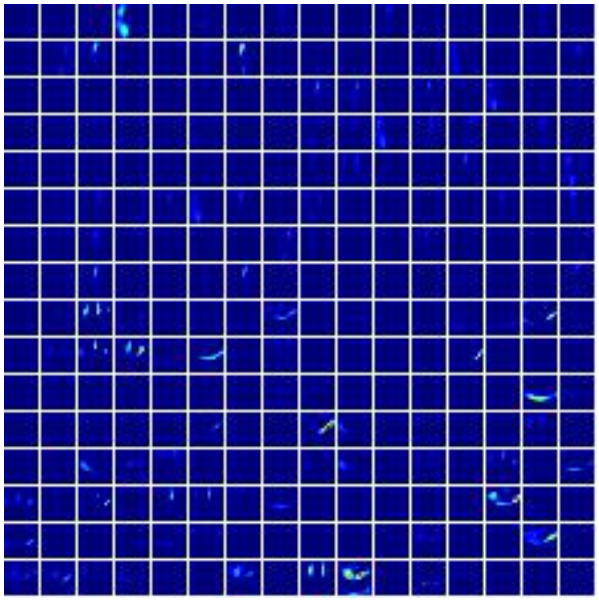
(c) Couche de convolution 2.



(d) Couche de convolution 3.

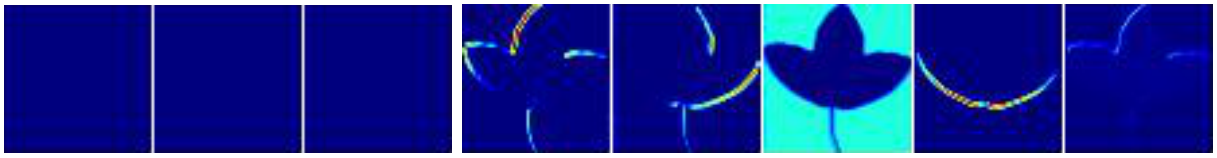


(e) Couche de convolution 4.



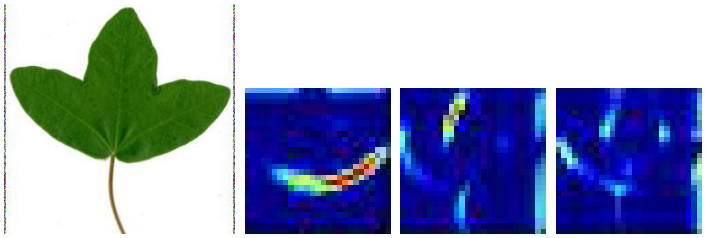
(f) Couche de convolution 5.

FIGURE 7.6 – Visualisation des images en sortie des noyaux de convolution des cinq couches de convolution du CNN AlexNet sur une image de feuille (*Acer monspessulanum*).

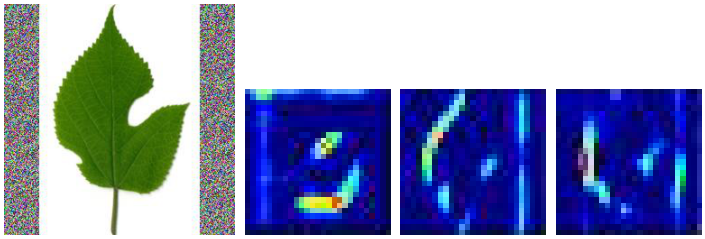


(a) Filtres non actifs. (b) Filtres actifs.

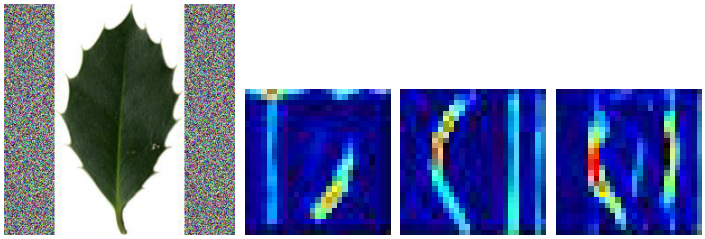
Figure 7.7 – Classification des filtres actifs et non actifs pour la première couche de convolution.



(a) Feuille de *Acer monspessulanum*.



(b) Feuille de *Broussonetia papyrifera*.



(c) Feuille de *Ilex aquifolium*.

Figure 7.8 – Filtres de la deuxième couche de convolutions séparant les contours des feuilles selon certaines orientations.

haut niveau.

Sur les couches de convolutions 3 et 4 (Figure 7.6), nous pouvons remarquer qu'il y a encore beaucoup de filtres qui s'activent pour faire ressortir les contours de la feuille. Nous pouvons distinguer un certain nombre de filtres s'activant sur la base de la feuille, ainsi que sur la pointe. En effet, la forme de l'apex et de la base permettent de distinguer les espèces d'arbres. Il semble donc que le réseau ait appris que ces éléments étaient pertinents pour l'identification des espèces végétales. Cependant, en regardant la réponse de certains filtres pour des feuilles possédant des caractéristiques communes, nous pouvons détecter des filtres spécialisés dans la détection ou non d'une caractéristique botanique. Par exemple sur la couche de convolutions numéro 3, nous avons repéré un filtre caractéristique de la forme de la base. Celui-ci s'active quand la base est large, et ne s'active plus quand la

base est fine (Figure 7.9).

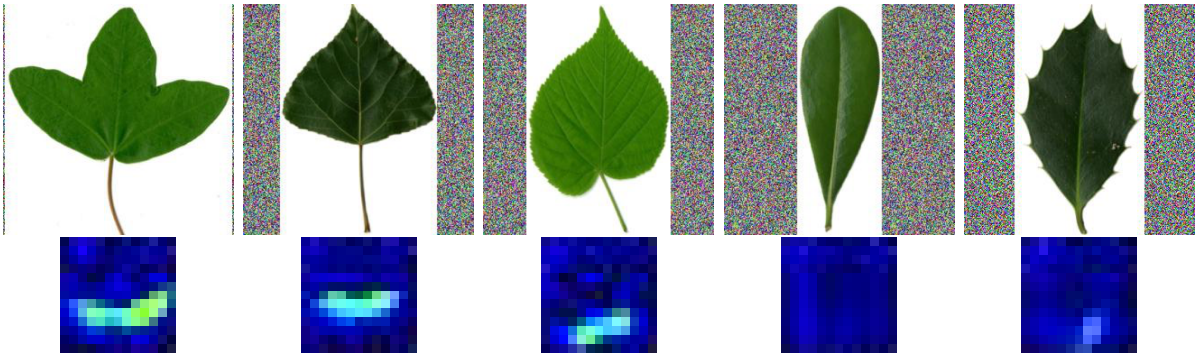


FIGURE 7.9 – Visualisation du même filtre de la couche de convolution 3 pour différentes feuilles avec des formes de base différentes.

Avoir un filtre qui ne s'active pas est aussi une forme d'indication de la forme d'une partie de la feuille. Dans l'exemple de la Figure 7.9, si le filtre repéré s'active, nous pouvons en déduire que la base est plutôt plate ou obtuse, tandis que s'il ne s'active pas, nous pouvons en déduire que la feuille a une base pointue. Des exemples semblables peuvent être trouvés parmi les filtres de la quatrième couche de convolution (Figure 7.10).

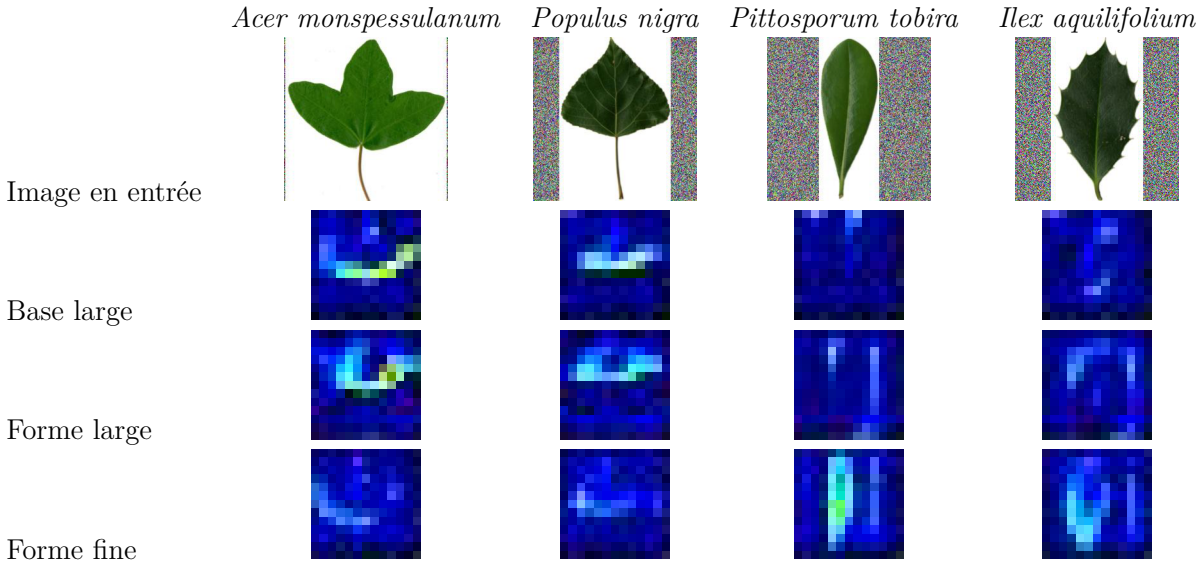


FIGURE 7.10 – Visualisation de filtres caractéristiques de la quatrième couche de convolution du réseau AlexNet.

De même que sur la couche de convolution numéro 3, il existe dans la couche de convolution numéro 4 un filtre qui est caractéristique de la forme de la base. Celui-ci est activé quand la base de la feuille est large, et inactivé quand la base est étroite. Nous avons aussi repéré un filtre qui est spécialisé pour caractériser la largeur de la feuille. Lorsque la feuille est large, par exemple pour *Acer monspessulanum* ou *Populus nigra*, le filtre est activé, et pour les feuilles de *Pittosporum tobira* ou *Ilex aquilifolium* le filtre est inactivé.

Il existe aussi un filtre qui est complémentaire au précédent, qui va s'activer quand la largeur de la feuille est faible, et qui se désactive quand la feuille est large (Figure 7.10).

La cinquième et dernière couche de convolution du CNN AlexNet est plus difficile à interpréter car le niveau d'activation d'une grande partie des filtres est assez faible (beaucoup de filtres sombre). Avec notre sélection de feuilles, nous avons cependant pu détecter un filtre caractéristique de la présence de lobes dans la feuille (Figure 7.11).

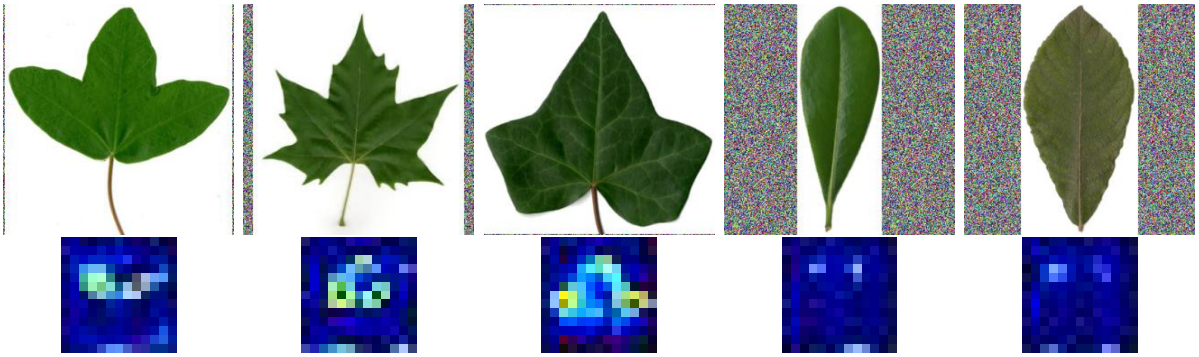


Figure 7.11 – Visualisation du même filtre de la couche de convolution 5 pour différentes feuilles lobées ou non.

Pour résumer, nous avons déduit à partir de la visualisation des filtres des couches de convolutions du CNN AlexNet que les deux premières couches extraient des informations de bas niveau sur les feuilles, c'est-à-dire les contours. Les contours sont extraits selon plusieurs orientations et à différents endroits de l'image, selon différentes intensités. C'est à partir de ces contours que les trois couches de convolution supérieures peuvent extraire des informations ayant un niveau sémantique plus élevé et pertinent pour la classification des espèces *via* les feuilles. En effet, nous avons pu observer que certains filtres retirent de l'information botanique, sur la forme générale, la base et l'apex.

Cependant, la détection des filtres caractéristiques a été réalisée "à la main", avec donc une interprétation pouvant être incomplète. Pour l'instant, nous n'avons été capable de détecter que les filtres donnant une information peu précise sur la forme de la feuille. Par exemple pour la base, nous avons trouvé un filtre indiquant si la feuille en entrée du réseau a une base large ou étroite, mais une base large peut correspondre à une base en cœur, tronquée, oblique, *etc.* Or, cette caractérisation plus précise doit être extraite par le réseau puisque celui-ci donne de bons taux de classification. Nous proposons donc d'automatiser le processus de détection de filtres actifs afin de faciliter leur détection lorsqu'une caractéristique botanique est présente sur une feuille.

7.3 Détection automatique des filtres actifs

Nous visualisons les filtres des trois dernières couches avec une image de la classe *Acer monspessulanum* (Figure 7.12). Les deux premières couches ne sont plus étudiées car

nous souhaitons nous intéresser aux informations de haut niveau portant des informations botaniques. La visualisation des filtres automatiquement détectés comme actifs va être réalisée sur la couche de convolution numéro 3. Les filtres vont être donnés en niveau de gris, et non plus avec le mappage de couleur "Jet".

FIGURE 7.12 – Feuille de *Acer monspessulanum* utilisée pour la visualisation des couches de convolutions.

7.3.1 Méthodes de traitement d'images basiques en vue de l'automatisation

Pour commencer, nous avons testé des méthodes basiques de traitement d'images afin de faire une séparation entre les filtres actifs et non actifs. Les résultats des filtres de convolutions de la troisième couche sont visibles sur la Figure 7.13.

Comme précédemment, nous pouvons remarquer que certains filtres sont activés et que d'autres non. Une méthode basique pour séparer les filtres actifs des filtres non actifs consiste à réaliser un seuillage. Pour chaque filtre, un seuillage est réalisé avec un seuil à 180. La valeur de seuil a été fixée empiriquement. Pour éviter la détection de filtres détectant les bords bruités des images, nous calculons le rapport entre le nombre de pixels supérieurs au seuil et le nombre de pixels total de l'image. Si ce rapport est supérieur ou égale à 1% le filtre est considéré comme actif. La Figure 7.14 présente le résultat de la binarisation et les filtres détectés comme actifs.

Sur la Figure 7.14b, les filtres encadrés en blanc sont les filtres détectés comme actifs par la méthode détaillée précédemment. Nous pouvons constater que les filtres très actifs (avec un nombre important de pixels possédant un niveau de gris élevé) ont été très bien détectés par la méthode. Cependant, un nombre notable de filtres ne sont pas considérés comme actifs (trois exemples sont encadrés en vert sur la Figure 7.14a). Cela provient du fait que le seuil de la binarisation et du rapport entre le nombre de pixels blancs et le nombre de pixels total du filtre sont fixés de manière empirique. Néanmoins si ces seuils

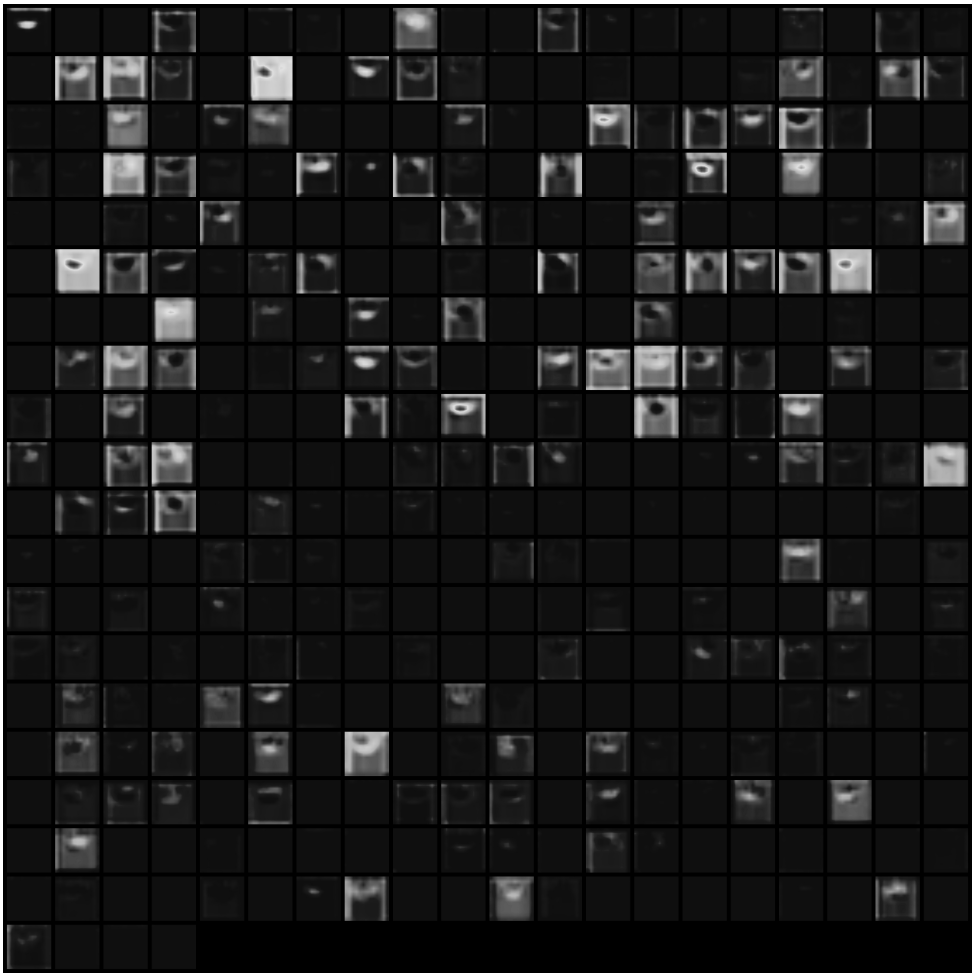


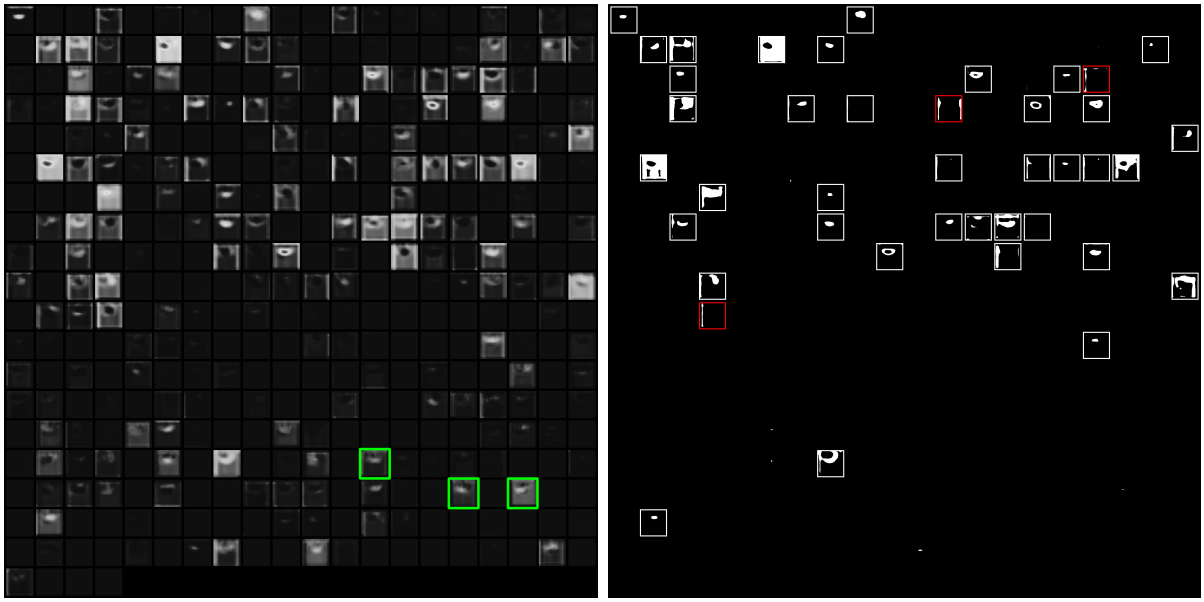
Figure 7.13 – Visualisation de la sortie des filtres de convolution de la couche numéro 3 du réseau AlexNet.

sont abaissés, des filtres mettant en valeur le bruit gaussien sur les bords de l'image sont déclarés comme actifs. Or, ils ne portent pas d'information sur la feuille. Malgré tout, certains filtres mettant en valeur les bords de l'image sont détectés comme des filtres actifs (filtres encadrés en rouge sur la Figure 7.14b).

Une deuxième méthode de détection mise en place est le calcul de la valeur moyenne de tous les pixels de chaque filtre. Pour déclarer le filtre comme actif ou non, un seuil a été de nouveau fixé. Si la valeur moyenne des pixels est supérieure à 40, le filtre correspondant est considéré comme actif. La Figure 7.15 présente encadrés en blanc les filtres classifiés comme actifs sur la troisième couche de convolutions.

Contrairement à la méthode précédente, cette méthode permet de relever plus de filtres qui nous semblent actifs. Toutefois, il reste encore des filtres (encadrés en rouge sur la Figure 7.15) qui ne sont pas détectés comme actifs, malgré la tache au centre. La prise en compte de l'écart-type de l'histogramme permettrait de pallier ce problème.

Une troisième méthode également testée consiste donc à calculer l'histogramme des niveaux de gris de chaque sortie de filtre pour ensuite en calculer l'écart-type. Nous



(a) Sortie originale de la couche 3. (b) Résultat de la binarisation des filtres.

FIGURE 7.14 – Détection des filtres de convolution actifs grâce à la méthode de binarisation.

considérons que si cet écart-type est compris dans l'intervalle [250 ; 270] alors le filtre est déclaré comme actif. La Figure 7.16 montre les filtres considérés comme actifs par cette troisième méthode.

Cette méthode simple donne de bons résultats malgré quelques erreurs de détection ou d'absence de détection. L'inconvénient majeur de cette méthode est le fait de fixer de manière empirique les limites de l'intervalle. Afin de pallier les différentes faiblesses des trois méthodes précédentes et de combiner leurs avantages, nous proposons une méthode un peu plus évoluée de détection de filtres actifs.

7.3.2 Méthode basée sur le seuillage d'Otsu

Pour rappel, nous considérons qu'un filtre est actif si celui si possède l'ensemble de ses pixels à un niveau de gris élevés, ou s'il y a une tâche importante par rapport au reste du filtre. De manière à détecter les filtres possédant ces caractéristiques, nous avons développé un algorithme avec plusieurs conditions. Premièrement, la sortie du filtre est seuillée avec un seuil à 75, et nous comptons le rapport N_1 de pixels supérieurs à ce seuil par rapport au nombre total de pixels en sortie du filtre. Si N_1 est supérieur à 95%, alors le filtre est considéré comme actif. Deuxièmement, si N_1 est compris entre 5 et 95%, le filtre est considéré comme potentiellement actif. Si N_1 est inférieur à 5%, le filtre est inactif. Si le filtre est considéré comme potentiellement actif, nous supposons que l'image en sortie du filtre n'est pas entièrement claire, mais contient une tache. Dans ce cas, son histogramme de niveaux de gris possède deux pics et nous pouvons utiliser la méthode de Otsu. Nous avons vu dans le Chapitre 2 que la méthode de Otsu [Ots79] permet de

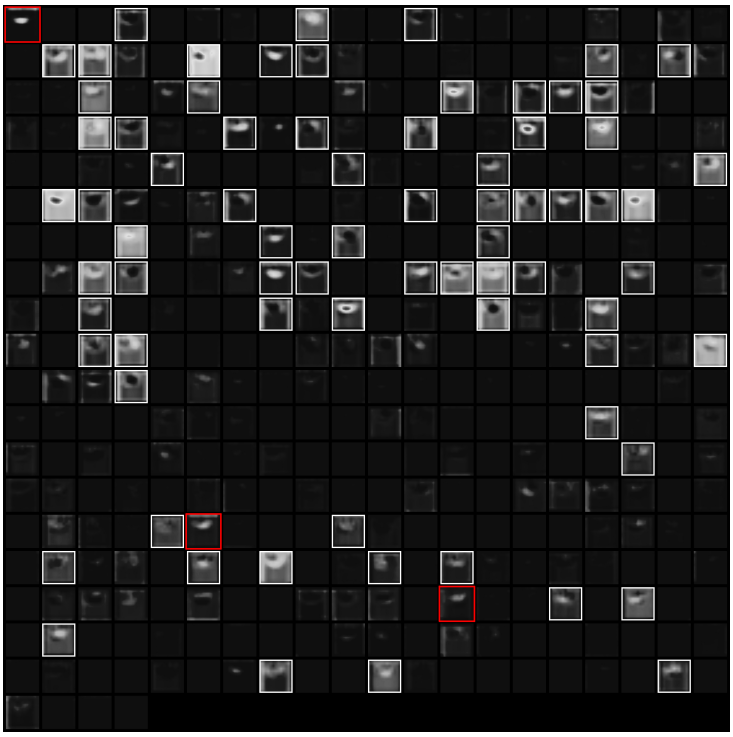


Figure 7.15 – Résultat de la détection de filtres actifs par calcul de la valeur moyenne des pixels.

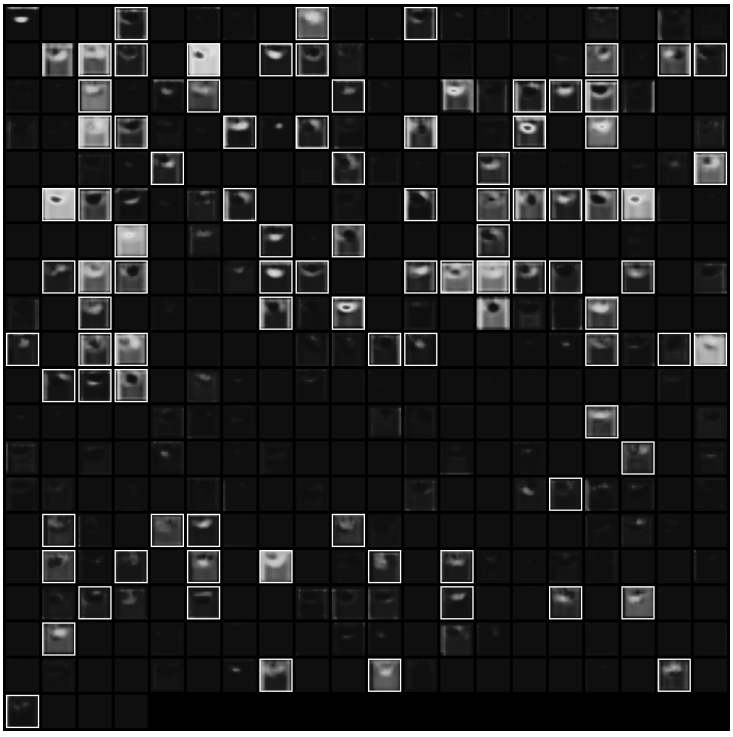


Figure 7.16 – Résultat de la détection de filtres actifs par calcul de l'écart-type de l'histogramme.

calculer automatiquement une valeur de seuil S optimale en fonction de l'histogramme de l'image. La sortie du filtre est ensuite seuillée avec ce seuil S. De nouveau, un rapport N_2 entre le nombre de de pixels blancs et le nombre de pixels total est calculé. Si N_2 est supérieur à 5%, alors le filtre est détecté comme actif.

Algorithm 3 Algorithme de calcul de la matrice de confusion.

```
int nbTestSample ← nombre d'images de test
Seuiller l'image à 75
Calculer le rapport N1
if N1 <5% then
  Filtre non actif
end if
if N1 >5% et N1 <95% then
  Utiliser l'algorithme d'Otsu pour trouverS
  Seuiller l'image à S
  Calculer le rapport N2
  if N2 >5% then
    Filtre actif
  end if
end if
if N1 >95% then
  Filtre actif
end if
```

La Figure 7.17 présente les filtres détectés comme actifs (encadrés en blanc) selon cette méthode.

Nous pouvons constater que cette méthode permet de détecter l'ensemble des filtres que nous considérons comme actifs, en palliant les inconvénients des méthodes précédentes. Il y a uniquement un seuil pour rejeter directement un filtre comme non actif (ensemble des pixels avec des niveaux de gris très faible) et un seuil pour valider directement si un filtre est actif (ensemble des pixels avec des niveaux de gris élevés). Dans le cas où il existe une incertitude, c'est-à-dire que la sortie du filtre présente une tache, le seuillage par la méthode d'Otsu permet d'automatiquement choisir le bon seuil pour faire la séparation entre le fond et l'élément pertinent mis en évidence par le filtre.

Si nous observons les filtres marqués comme actifs et encadrés en rouge, nous pouvons remarquer qu'ils ne semblent pas mettre en évidence une partie de la feuille. En l'occurrence, ils mettent en évidence les bords bruités de l'image. Il s'avère que le réseau de neurones AlexNet est sensible à la méthode de rajout de bruit sur les bords de l'image afin d'obtenir des images de taille carrée sans perdre le ratio de la feuille.

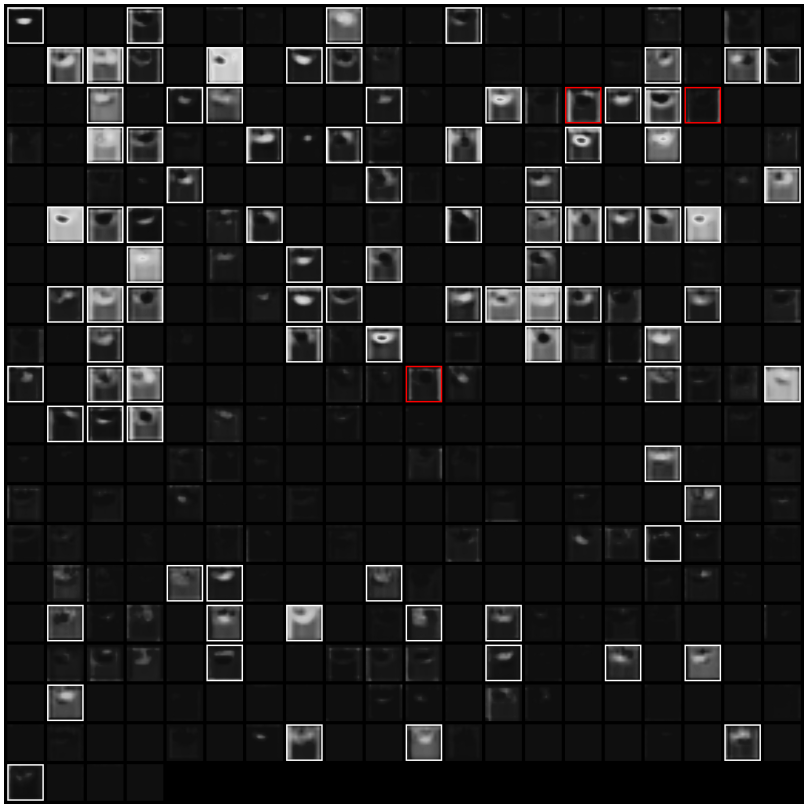


Figure 7.17 – Résultat de détection des filtres actifs avec la méthode basée sur le seuillage d'Otsu.

7.4 Identification des filtres révélateurs d'une caractéristique botanique

À présent que les filtres actifs sont méthodiquement détectés, nous souhaitons détecter automatiquement à quelle caractéristique botanique ceux-ci sont associés. Par exemple dans la partie 7.2.2, nous avions identifié un filtre caractéristique de la forme de la base (Figure 7.9). D'autre part, nous avons vu dans Chapitre 1 que chaque partie de la feuille (forme générale, apex, base, marge) a sa forme définie par un vocabulaire botanique.

Pour chaque espèce de notre base de données, nous avons accès aux formes possibles que peuvent prendre les différentes parties. Par exemple, les feuilles de l'espèce *Acer monspessulanum* ont une forme de limbe palmatilobée, une base en cœur, un sommet de forme subaiguë ou obtuse, et une marge sinuée ou lobée. Ainsi, pour chaque composante de la feuille, nous pouvons connaître toutes les espèces qui ont des feuilles avec une caractéristique locale précise.

Afin de détecter automatiquement les filtres caractéristiques d'une forme de composante de la feuille, toutes des images de feuilles des espèces ayant cette forme (par exemple base tronquée) sont données en entrée du réseau et toutes les sorties de filtres des couches 3, 4 et 5 sont sauvegardées. Par exemple, pour détecter si un filtre est caractéristique de la base tronquée, toutes les feuilles des espèces *Alnus glutinosa*, *Quercus coccifera*, *Quercus*

pubescens et *Sorbus torminalis*, qui ont une base tronquée, sont données en entrée du réseau (voir Figure 7.18).

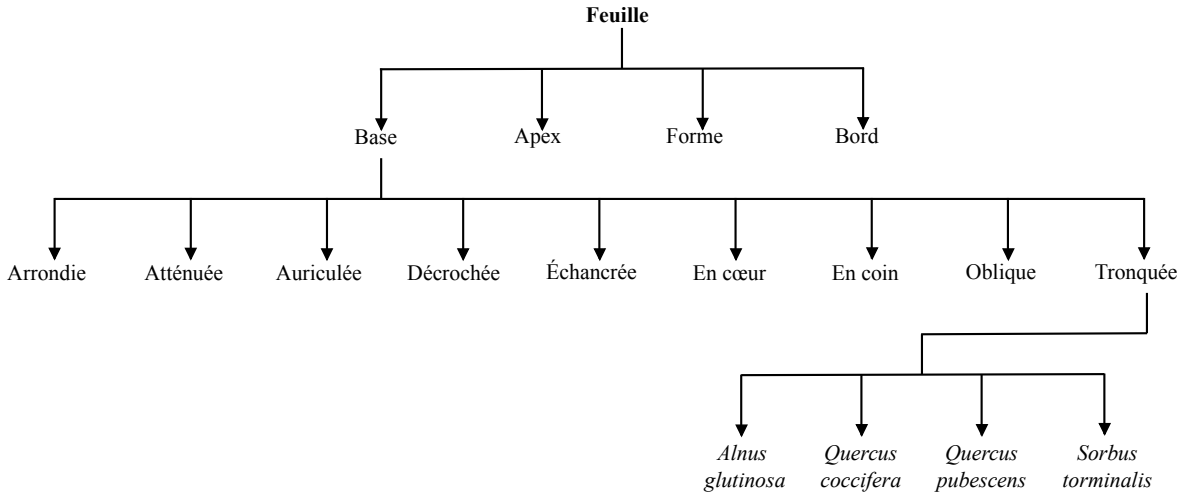


FIGURE 7.18 – Exemple d'espèces avec une caractéristique de base tronquée.

Une fois tous les filtres d'une couche récupérés, pour toutes les feuilles ayant une base tronquée, nous appliquons notre méthode de détection de filtre actif basée Otsu. Nous calculons ensuite pour chaque filtre le nombre de fois que celui-ci est activé. Ce nombre est normalisé par le nombre d'image dans chaque classe. Afin d'avoir une représentation visuelle, les valeurs normalisées du nombre d'activations sont multipliées par 255, afin d'obtenir une image en niveau de gris. La Figure 7.19 présente le niveau d'activation de chaque filtre pour les feuilles ayant une base tronquée sur la troisième couche de convolution du réseau AlexNet.

Plus le filtre présente un niveau de gris élevé, plus celui-ci s'est activé au passage des différentes images, tandis qu'un filtre sombre indique que celui-ci ne s'active que très peu. Nous pouvons constater que de nombreux filtres s'activent régulièrement. Cependant, nous ne pouvons garantir que ces filtres sont caractéristiques de la base tronquée. En effet, il se peut que toutes les feuilles qui ont une base tronquée possèdent aussi une marge doublement dentée par exemple. Nous pouvons voir sur la Figure 7.20 les niveaux d'activation que nous obtenons pour des feuilles avec une marge doublement dentée ou un apex aigu.

Nous remarquons que ces visualisations des niveaux d'activation de filtres sont très semblables. Ceci montre que les filtres qui s'activent régulièrement ne sont pas uniquement caractéristiques d'une seule forme d'une partie de la feuille. De manière à identifier quel est le, ou quels sont les, filtres caractéristiques de ces formes, nous calculons les niveaux d'activation des filtres pour toutes les feuilles, afin d'avoir une valeur moyenne. Ce niveau moyen d'activation des filtres est pondéré par la population de feuilles possédant une forme définie. Par exemple, pour avoir le niveau moyen d'activation des filtres pour la base, nous

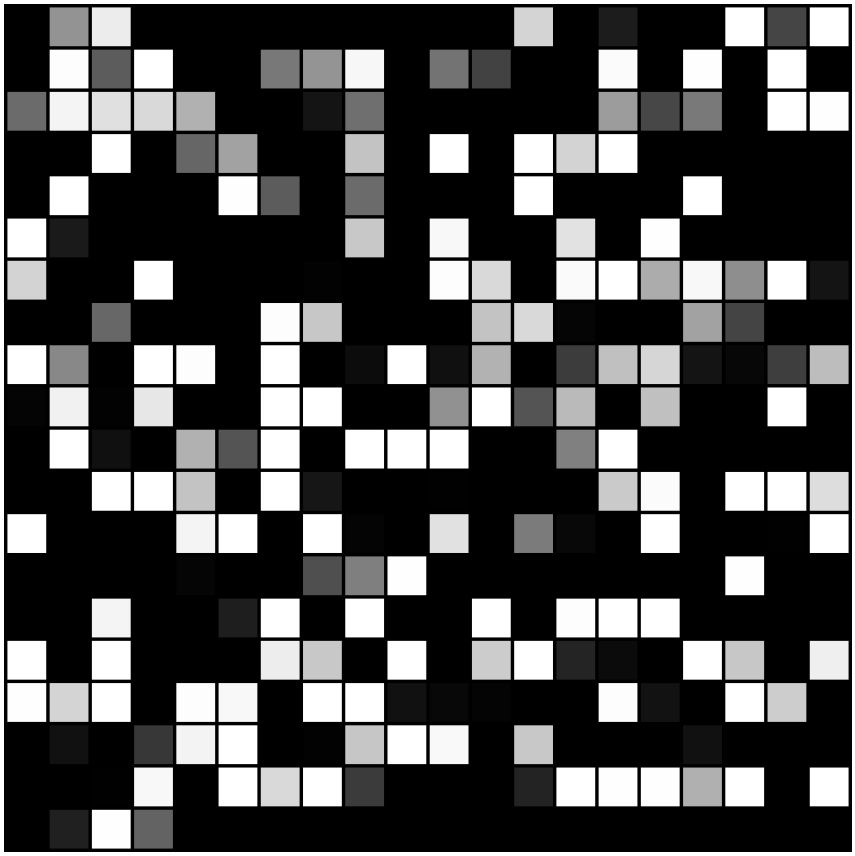
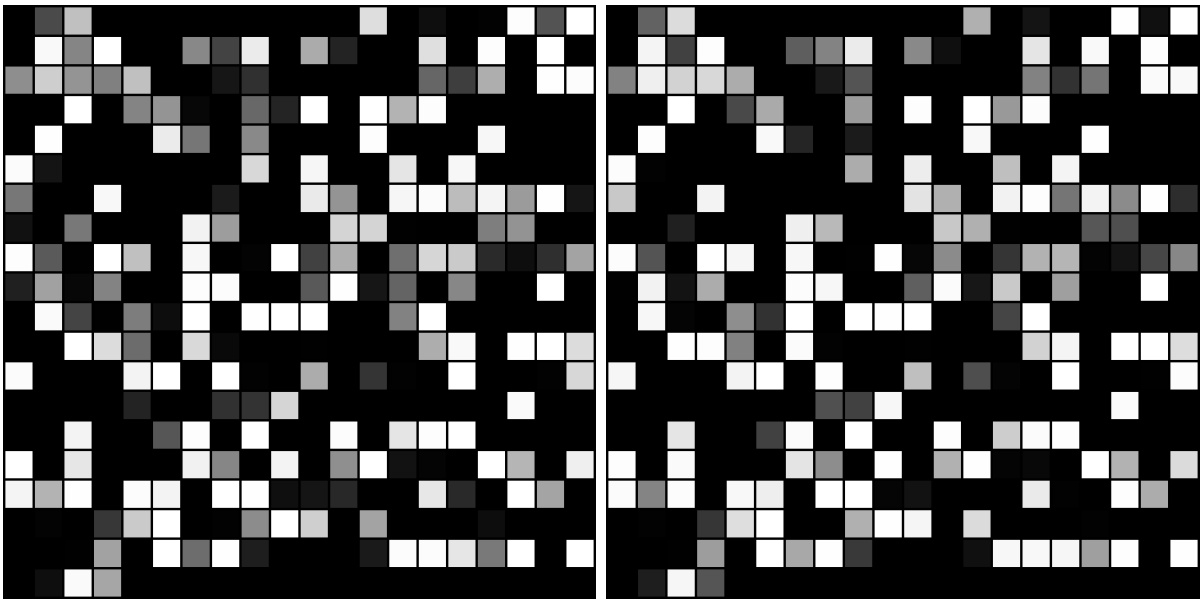


Figure 7.19 – Niveau d'activation des filtres de la troisième couche de convolution pour des feuilles présentant une base tronquée.

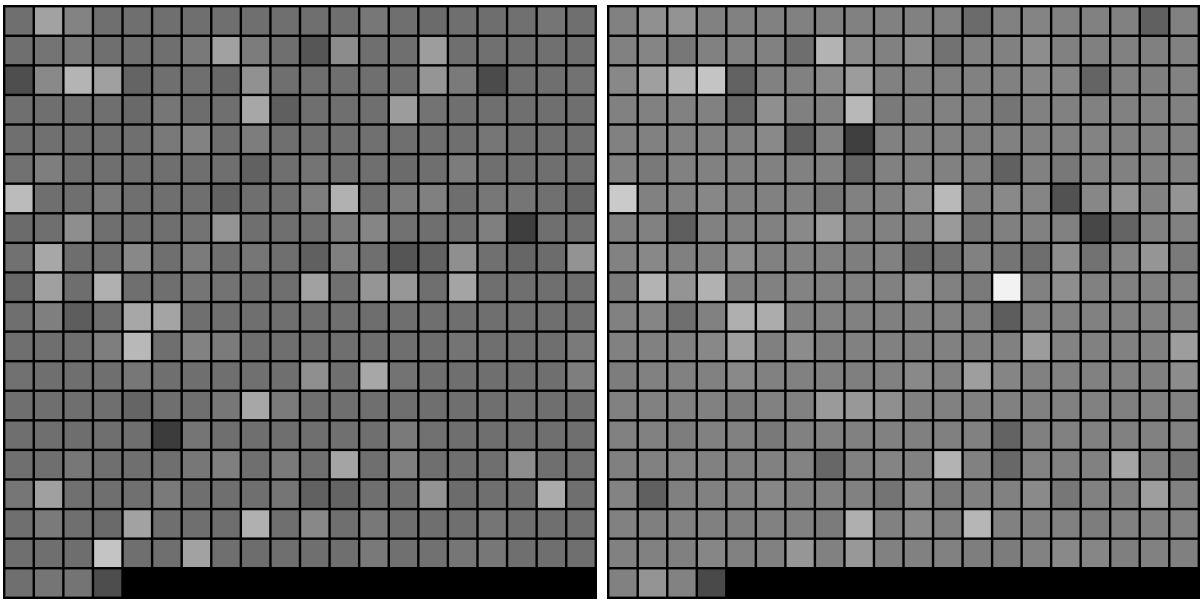


(a) Apex aigu. (b) Marge doublement dentée.

Figure 7.20 – Autres exemples de niveaux d'activation pour la troisième couche de convolution.

calculons les niveaux d'activation pour les feuilles ayant une base arrondie, puis atténuée, puis auriculée, *etc.* Ces niveaux d'activation vont ensuite être pondérés par le nombre de

feuilles dans chaque classe de base. Enfin, en soustrayant le nombre d'activation de chaque filtre par la valeur moyenne, nous pouvons obtenir les filtres caractéristiques d'une forme de composante de la feuille. La figure 7.21 montre les filtres caractéristiques de la base tronquée et du bord doublement denté après soustraction de la moyenne.



(a) Base tronquée. (b) Marge doublement dentée.

FIGURE 7.21 – Activation des filtres de la troisième couche de convolution pour la base tronquée après soustraction de la moyenne.

À présent les niveaux d'activation ne sont plus aussi semblables pour des caractéristiques différentes. Grâce à la soustraction de la moyenne, nous pouvons plus facilement détecter les filtres caractéristiques d'une forme d'un élément de la feuille. Si nous étudions la Figure 7.21b, nous pouvons visualiser un filtre avec un niveau de gris élevé se situant colonne 13 ligne 9 (en commençant par colonne 0 et ligne 0 en haut à gauche). Ainsi nous pouvons affirmer que ce filtre s'active quand la feuille donnée en entrée du réseau de neurones AlexNet présente une marge doublement dentée. Avoir un filtre avec un niveau d'activation plus élevé que la moyenne (niveau de gris plus élevé) signifie que celui-ci s'active au passage d'image de feuille présentant une caractéristique particulière. Mais à l'opposé, si un filtre présente un niveau d'activation plus faible que la moyenne (niveau de gris faible), cela signifie que ce filtre ne s'active pas au passage d'une image de feuille ne présentant pas cette caractéristique particulière. Ces filtres nous apportent donc aussi une information sur la forme des caractéristiques botaniques de la feuille.

7.5 Conclusion

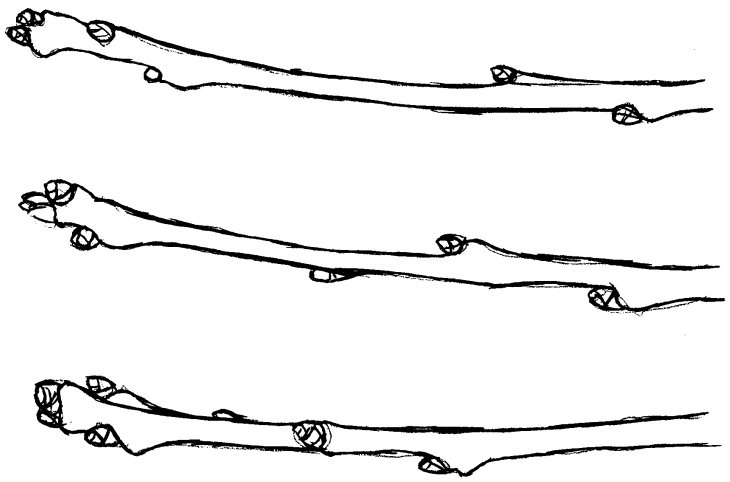
En premier lieu, nous avons vu dans ce chapitre qu'il était possible d'utiliser un réseau de neurones à convolutions assez basique (AlexNet, 5 couches de convolutions) pour

classifier une feuille d'arbre photographiée sur un fond uni parmi 72 espèces. En outre, nous avons vu grâce à la visualisation des sorties des filtres des différentes couches de convolution que le CNN AlexNet extrait des caractéristiques de bas niveaux, telles que les contours, sur les deux premières couches et des caractéristiques de haut niveau sur les trois dernières couches. Ces caractéristiques de haut niveau sont des caractéristiques botaniques, qui viennent définir la forme générale du limbe, la forme de la base et de l'apex.

La détection manuelle des filtres révélateurs de certaines caractéristiques botaniques est chronophage et peu précise. C'est pourquoi nous avons proposé de détecter automatiquement les filtres considérés comme actifs grâce à des méthodes de traitement d'images. Ces méthodes sont des outils assez simples mais qui donnent de bons résultats. Grâce à la détection automatique de ces filtres, nous avons pu mettre en place une méthodologie pour identifier cette fois-ci les filtres spécifiques à une seule caractéristique botanique, telle qu'une base tronquée ou une marge doublement dentée. Notre méthode permet de détecter une caractéristique botanique sur une feuille par l'activation ou la non-activation d'un filtre particulier.

Une des contraintes que nous avons cependant relevée est la mise en évidence des bords bruités des images par les filtres de convolution, entraînant leur détection en tant que filtre actif. Une des solutions serait peut-être de prendre les couleurs des bords de l'image et les dupliquer là où il manque de l'information. Nous ne pouvons pas directement remplir avec des pixels blancs, car toutes les feuilles ne sont pas photographiées sur une feuille blanche (Figure 7.2). Ce travail reste un premier pas vers une exploitation future des CNN comme outil de reconnaissance d'éléments botaniques, afin de prédire l'espèce végétales à partir de la classification de la forme du limbe, de la forme de l'apex, *etc.*

Conclusion générale



Tout au long de ce manuscrit, nous avons vu comment il était possible de retirer des informations des différents organes de l'arbre que sont les feuilles, les écorces, les fleurs et les fruits, ainsi que la combinaison de ceux-ci, en vue d'une identification de l'espèce végétale. Ce processus reste tout de même complexe du fait, entre autres, de la nature vivante du sujet. Malgré les différentes contraintes, nous avons apporté un certain nombre de contributions.

Segmentation des organes végétaux

Lors du projet ReVeS, une méthode très efficace de segmentation de feuille en milieu naturel a été proposée [Cer13b], nommée *GAC*. Cependant, cette méthode n'a pas pu être transposée sur les autres types d'organes à cause de la très grande variabilité de forme de ceux-ci. Nous avons donc proposé de remplacer le modèle de forme qui contraint la segmentation pour une carte de profondeurs qui contraint la segmentation à être réalisée au premier plan. La prise en compte de la profondeur a été rendue possible grâce à l'utilisation des doubles caméras en face arrière qui équipent de plus en plus les smartphones. Par ailleurs, une carte de dissimilarité de couleurs et de textures a été proposée afin de segmenter l'objet d'intérêt situé dans le premier plan grâce à un MLP de très faible dimension. La segmentation *via* contour actif sur cette carte a également montré son efficacité sur une base de données contenant d'autres objets que des fruits et des fleurs. Cette segmentation requiert une petite implication de la part de l'utilisateur pour dessiner une trace sur l'objet d'intérêt qu'il souhaite segmenter. Une grande importance est donnée à cette trace étant donné qu'elle permet de construire les deux cartes ainsi que de fixer leurs poids respectifs lors de la segmentation. L'influence de cette trace a été étudiée, et elle peut faire échouer le processus de segmentation si celle-ci est négligée. Concernant l'écorce, nous avons fait le choix de ne pas mettre en place un algorithme de segmentation car nous supposons que nous pourrons indiquer à l'utilisateur comment photographier le tronc *via* l'application pour que celui-ci ne prenne en photo que l'écorce en se rapprochant suffisamment de l'arbre.

Une des perspectives serait de mettre en place un système de segmentation pour les écorces en fonction des retours de l'utilisation de l'application d'aide à l'identification des espèces d'arbres. En effet, si nous nous rendons compte que la demande précédente est incomprise ou irréalisable au risque d'avoir une image floue, nous pourrons intégrer un outil pour dessiner une ROI rectangulaire sur l'image ou mettre en place un algorithme de segmentation de texture. D'autre part, la segmentation des fruits et des fleurs est réalisée en un temps raisonnable sur ordinateur. Il sera nécessaire de vérifier que ces temps restent acceptables quand l'algorithme sera implémenté sur un smartphone.

Description des organes végétaux

Une des lignes directrices de cette thèse est la vocation pédagogique de l'application. Ainsi les descripteurs proposés ont été créés, ou choisis, de manière à intégrer des informations sémantiques et botaniques. Concernant les écorces, nous avons proposé un descripteur (mot) permettant de caractériser la texture verticale et horizontale des écorces. De plus, nous avons intégré l'information de couleur à la reconnaissance des écorces. Une autre contrainte de cette thèse est l'utilisation des algorithmes sur smartphone. Ainsi les descripteurs se doivent d'être les plus petits possible. Afin de réduire notre descripteur de texture, nous avons proposé une fusion des bins des couleurs les moins représentées dans les écorces. Cette méthode permet de réduire significativement le vecteur de couleur de 180 bins à seulement 30. En ce qui concerne les fleurs et les fruits, nous n'avons pu créer des vecteurs descriptifs des organes s'appuyant sur les caractéristiques botaniques au vu de la très grande variabilité de ceux-ci et du fait que nous traitons indifféremment les fruits et les fleurs, car la distinction entre ces organes est complexe pour un novice en botanique. Les descripteurs choisis caractérisent cependant la forme, la couleur et la texture : caractéristiques cependant utilisées, entre autres, dans les ouvrages botaniques.

Nous pensons que le descripteur de textures d'écorces que nous avons proposé dans ce manuscrit pourra être utilisé pour caractériser la texture sur d'autres bases de données telles que Outex [2] par exemple. D'autre part, nous avons montré que la réduction de l'histogramme colorimétrique n'était pas utile dans le cadre de la reconnaissance de fruits et de fleurs, mais elle pourrait l'être pour un autre jeu de données, et pas uniquement sur un histogramme de couleurs. Enfin, pour l'instant, seules la texture, la couleur et la forme qui décrivent les fruits et les fleurs. La classification de ces organes pourrait être améliorée si d'autres descripteurs étaient ajoutés. Nous pourrions rajouter, par exemple, un descripteur pour révéler la présence de ronds, traduisant la présence de baies dans l'objet d'intérêt.

Classification et combinaison

Selon les différents chapitres de ce manuscrit, nous avons vu que chaque organe avait un classifieur différent. Le classifieur pour les feuilles, implémenté dans l'application *Folia* n'a pas été modifié car donnant de très bons résultats. Pour les écorces, le meilleur classifieur pour les descripteurs s'est révélé être un SVM un-contre-tous. Cet ensemble de SVM binaires permet d'ordonner les prédictions d'appartenance aux différentes espèces tout en obtenant de meilleurs résultats de classification qu'un réseau de neurones (MLP). Cependant c'est l'inverse pour la classification des fruits et des fleurs. En effet, c'est un MLP qui est utilisé pour réaliser la classification de ces organes et qui donne de meilleurs

2. http ://www.outex.oulu.fi/

taux qu'avec l'utilisation d'un SVM. Pour les fruits et les fleurs, nous avons aussi considéré leur nature botanique, c'est-à-dire leur similitude traduisant un rapprochement des espèces dans la classification botanique (genre et famille). Grâce aux matrices de confusion, nous avons montré qu'il existait en effet de la confusion entre les fruits et les fleurs des espèces d'arbres appartenant à un même genre, et entre les fleurs des espèces d'arbres appartenant à une même famille.

D'autre part, nous avons utilisé cet outil qu'est la matrice de confusion afin de raffiner nos résultats de classification. Nous avons montré qu'à partir de données d'apprentissage il était possible de prendre en compte la confusion qui existe entre les espèces, par exemple à cause de la similarité inter espèces pour les écorces. Ceci permet d'augmenter les bons taux de classification des espèces, notamment aussi en palliant les problèmes de déséquilibre des populations des classes. Nous avons bon espoir que la prise en compte des matrices de confusion améliore les performances d'autres systèmes de classification dans d'autres applications.

Il ne suffit pas de classifier les organes un à un indépendamment les uns des autres. Dans ce manuscrit, nous avons ainsi proposé une méthode de combinaison *a posteriori* des différents résultats de classification des différents organes. Nous avons vérifié que mixer les prédictions des différents organes n'apportait pas de la confusion. Effectivement, combiner deux ou trois organes augmente systématiquement les taux de classification par rapport à un seul organe considéré. Cependant la méthode proposée reste relativement basique, et ne prend pas vraiment en compte la nature botanique de notre objet d'étude. En effet, considérer uniquement la prédiction du genre apportée par la classification des fruits-fleurs ne permet pas d'augmenter les résultats de classification des espèces par rapport à l'information directement de l'espèce. Nous pouvons alors espérer trouver une meilleure méthode de combinaison pour prendre en compte les clades botaniques, genre ou famille, par exemple en partenariat avec nos collègues de projet du LISTIC, spécialisés en fusion de données.

Deep Learning

Pour finir, nous avons tenu dans ce manuscrit à intégrer des comparaisons avec le Deep Learning au vu de sa grande utilisation dans le domaine du traitement d'images. Ainsi, nous avons pu remarquer que les réseaux de neurones à convolution donnent de meilleurs résultats que notre descripteur sur les écorces grâce au nombre suffisamment important de patchs que nous avons pu extraire sur les images de notre base de données. Cependant, les résultats des CNN sont inférieurs aux taux de classification que nous obtenons avec notre descripteur pour la reconnaissance des fruits et des fleurs, malgré une augmentation du nombre de données.

Toutefois, nous avons pu explorer les différentes couches de convolution d'un CNN

afin que celui-ci ne fonctionne pas uniquement tel une "boîte noire". Ainsi, nous avons découvert que sur les dernières couches de convolutions des informations botaniques étaient extraites. En mettant en place une méthode de détection de filtres actifs, nous avons pu automatiser la localisation de filtres révélateurs d'une caractéristique botanique telle que le bord doublement denté, ou une base tronquée. Ainsi, à partir d'un CNN au préalablement entraîné sur une base de données de feuilles, nous pouvons indiquer à l'utilisateur si la feuille qu'il a prise en photo a une base tronquée ou non, *etc.* Une perspective de travail est donc d'étudier, d'une part, la portabilité d'un CNN sur un smartphone, et d'autre part, la possibilité d'extraire des filtres toutes les caractéristiques botaniques.

Perspectives pour l'identification des végétaux

Les travaux présentés dans ce manuscrit représentent une base solide pour l'extension de l'application d'aide à l'identification des espèces d'arbres *Folia*. Son application sur smartphone reste à être évaluée, ainsi que le retour des utilisateurs concernant les différentes interactions avec l'application. Dans la suite du projet ReVeRIES, les algorithmes de traitement d'images pour l'identification des végétaux ont pour but d'être utilisés en tant qu'"expert botanique". Afin d'enseigner la botanique, l'application demandera à l'utilisateur de photographier un gland, ou de trouver une feuille de platane. *Folia* viendra ensuite valider, ou non, les réponses de l'utilisateur. Le fait que nos descripteurs se basent sur des caractéristiques botaniques sert cet objectif.

Enfin pour l'instant les informations servant à l'identification des espèces d'arbres sont uniquement extraites sur les différents organes de la plante, c'est-à-dire la feuille, l'écorce, les fleurs et les fruits. Cependant d'autres indices peuvent rentrer en compte dans le processus d'identification. Actuellement, dans l'application *Folia*, c'est la position GPS qui peut être prise en compte afin de réorganiser la liste d'espèces proposées à l'utilisateur. En effet, il est plus ou moins probable de trouver certaines espèces dans une région de France (Figure 7.22) en fonction du climat (montagnard, méditerranéen, *etc.*), de l'ensoleillement ou de l'altitude.

(a) Chêne rouvre (*Quercus petraea*)

(b) Érable de Montpellier (*Acer monspessulanum*)

FIGURE 7.22 – Cartes de plausibilité obtenues à partir des présences théoriques des espèces (0 en blanc et 1 en vert foncé). Illustrations issues de [Cer13b].

De même que pour les coordonnées GPS de l'arbre, des informations importantes à considérer seraient les périodes de floraison et de fructification des arbres. En fonction de la date de prise de photo de la fleur ou du fruit, nous pourrions raffiner la prédiction des espèces en fonction de la probabilité de floraison et de fructification à cette période.

Bibliographie

[Aff+17] Antoine Affouard et al. « Pl@ ntNet app in the era of deep learning ». In : *ICLR 2017 Workshop Track-5th International Conference on Learning Representations*. 2017.

[Ari+10] S Arivazhagan et al. « Fruit recognition using color and texture features ». In : *Journal of Emerging Trends in Computing and Information Sciences* **1**.*2* (2010), page(s): 90–94.

[Aro+12] Akhil Arora et al. « A Plant Identification System using Shape and Morphological Features on Segmented Leaflets : Team IITK, CLEF 2012. » In : *CLEF*. 2012.

[Bak+13] Vera Bakić et al. « Inria's participation at imageclef 2013 plant identification task ». In : *CLEF (Online Working Notes/Labs/Workshop) 2013*. 2013.

[BTV06] Herbert Bay, Tinne Tuytelaars et Luc Van Gool. « Surf : Speeded up robust features ». In : *European conference on computer vision*. Springer. 2006, page(s): 404–417.

[BCT17a] Sarah Bertrand, Guillaume Cerutti et Laure Tougne. « Bark Recognition to Improve Leaf-based Classification in Didactic Tree Species Identification ». In : *VISAPP 2017-12th International Conference on Computer Vision Theory and Applications*. 2017.

[BCT17b] Sarah Bertrand, Guillaume Cerutti et Laure Tougne. « Segmentation Algorithm on Smartphone Dual Camera : Application to Plant Organs in the Wild ». In : *ICMV 2017 - The 10th International Conference on Machine Vision*. 2017.

[BCT18a] Sarah Bertrand, Guillaume Cerutti et Laure Tougne. « Segmentation algorithm on smartphone dual camera : application to plant organs in the wild ». In : *Tenth International Conference on Machine Vision (ICMV 2017)*. T. 10696. International Society for Optics et Photonics. 2018, page(s): 106960T.

[BCT18b] Sarah Bertrand, Guillaume Cerutti et Laure Tougne. « Visualization of Leaf Botanical Features Extracted from AlexNet Convolutional Layers ». In : *IAMPS 2018 - 6th edition of the International Workshop on Image Analysis Methods for the Plant Sciences*. 2018.

[Ber+18] Sarah Bertrand et al. « Bark and leaf fusion systems to improve automatic tree species recognition ». In : *Ecological Informatics* (2018).

[BM92] Serge Beucher et Fernand Meyer. « The morphological approach to segmentation : the watershed transformation ». In : *Optical Engineering-New York-Marcel Dekker Incorporated-* **34** (1992), page(s): 433–433.

[BKK13] Anant Bhardwaj, Manpreet Kaur et Anupam Kumar. « Recognition of plants by leaf image using moment invariant and texture analysis ». In : *International Journal of Innovation and Applied Studies* **3**.*1* (2013), page(s): 237–248.

[Bon08] Georges Bonnier Gaston et De Layens. *Flore complète de la France et de la Suisse*. Librairie Generale, 1908.

[BYB15] Safia Boudra, Itheri Yahiaoui et Ali Behloul. « A Comparison of Multi-scale Local Binary Pattern Variants for Bark Image Retrieval ». In : *Advanced Concepts for Intelligent Vision Systems*. Springer. 2015, page(s): 764–775.

[Can87] John Canny. « A computational approach to edge detection ». In : *Readings in Computer Vision*. Elsevier, 1987, page(s): 184–203.

[Cas+12] Dalcimar Casanova et al. « IFSC/USP at ImageCLEF 2012 : Plant Identification Task. » In : *CLEF (Online Working Notes/Labs/Workshop)*. 2012.

[Cer13a] Guillaume Cerutti. « Segmentation and interpretation of natural images for tree leaf identification on smartphones ». working paper or preprint. Nov. 2013. url : `https://hal.archives-ouvertes.fr/hal-01461461`.

[Cer13b] Guillaume Cerutti. « Segmentation et interprétation d'images naturelles pour l'identification de feuilles d'arbres sur smartphone ». Thèse de doct. Université Lumière Lyon 2, 2013.

[Cer+11] Guillaume Cerutti et al. « Guiding active contours for tree leaf segmentation and identification ». In : *CLEF 2011, Conference on Multilingual and Multimodal Information Access Evaluation*. 2011, page(s): 1.

[Cer+12] Guillaume Cerutti et al. « Reves participation-tree species classification using random forests and botanical features ». In : *Conference and Labs of the Evaluation Forum (CLEF)*. 2012, page(s): 1.

[Cer+13a] Guillaume Cerutti et al. « Late information fusion for multi-modality plant species identification ». In : *Conference and Labs of the Evaluation Forum*. 2013, page(s): Working–Notes.

[Cer+13b] Guillaume Cerutti et al. « Understanding leaves in natural images–a model-based approach for tree species identification ». In : *Computer Vision and Image Understanding* **117**.*10* (2013), page(s): 1482–1501.

[CLZ11] Yuning Chai, Victor Lempitsky et Andrew Zisserman. « Bicos : A bi-level co-segmentation method for image classification ». In : *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE. 2011, page(s): 2579–2586.

[CPB15] Jyotismita Chaki, Ranjan Parekh et Samar Bhattacharya. « Plant leaf recognition using texture and shape features with neural classifiers ». In : *Pattern Recognition Letters* **58** (2015), page(s): 61–68.

[Cha+15] Julien Champ et al. « A comparative study of fine-grained classification methods in the context of the lifeclef plant identification challenge 2015 ». In : *CLEF : Conference and Labs of the Evaluation forum*. T. 1391. 2015.

[CHC03] Zheru Chi, Li Houqiang et Wang Chao. « Plant species recognition based on bark patterns using novel Gabor filter banks ». In : *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*. T. 2. IEEE. 2003, page(s): 1035–1038.

[Cho15] Sungbin Choi. « Plant identification with deep convolutional neural network : Snumedinfo at lifeclef plant identification task 2015 ». In : *Working notes of CLEF 2015 conference*. 2015.

[Cos03] Charles Coste Hippolyte et Flahault. *Flore descriptive et illustrée de la France, de la Corse, et des contrées limitrophes*. T. 2. P. Klincksieck, 1903.

[Don+15] Xingping Dong et al. « Interactive Cosegmentation Using Global and Local Energy Optimization. » In : *IEEE Trans. Image Processing* **24**.*11* (2015), page(s): 3966–3977.

[DJ12] Shiv Ram Dubey et Anand Singh Jalal. « Detection and classification of apple fruit diseases using complete local binary patterns ». In : *Computer and Communication Technology (ICCCT), 2012 Third International Conference on*. IEEE. 2012, page(s): 346–351.

[DJ15] Shiv Ram Dubey et Anand Singh Jalal. « Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning ». In : *International Journal of Applied Pattern Recognition* **2**.*2* (2015), page(s): 160–181.

[Eis17] Wilhelm Eisenreich. *Guide de la faune et de la flore*. Flammarion, 2017.

[Far03] Gunnar Farnebäck. « Two-frame motion estimation based on polynomial expansion ». In : *Scandinavian conference on Image analysis*. Springer. 2003, page(s): 363–370.

[Gar+15] Farid Garcia-Lamont et al. « Color characterization comparison for machine vision-based fruit recognition ». In : *International Conference on Intelligent Computing*. Springer. 2015, page(s): 258–270.

[GMC15] Z Ge, Chris Mccool et P Corke. « Content specific feature learning for fine-grained plant classification ». In : *Working notes of CLEF 2015 conference*. 2015.

[God12] Jean-denis Godet. *Guide des écorces des arbres d'Europe - Reconnaître et comparer les espèces*. Delachaux, 2012.

[GBJ16] Hervé Goëau, Pierre Bonnet et Alexis Joly. « Plant identification in an open-world (lifeclef 2016) ». In : *CLEF 2016-Conference and Labs of the Evaluation forum*. 2016, page(s): 428–439.

[GJB15] Hervé Goëau, Alexis Joly et Pierre Bonnet. « LifeCLEF Plant Identification Task 2015 ». In : *CLEF working notes 2015*. 2015.

[Goë+13] Hervé Goëau et al. « The imageclef 2013 plant identification task ». In : *CLEF*. 2013.

[Goë+14] Hervé Goëau et al. « Lifeclef plant identification task 2014 ». In : *CLEF2014 Working Notes. Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014*. CEUR-WS. 2014, page(s): 598–615.

[Gra+15] Manuel Grand-Brochier et al. « Tree Leaves Extraction in Natural Images : Comparative Study of Preprocessing Tools and Segmentation Methods ». In : *IEEE Transactions on Image Processing* **24**.*5* (2015), page(s): 1549–1560.

[GSM10] DS Guru, YH Sharath et S Manjunath. « Texture features and KNN in classification of flower images ». In : *IJCA, Special Issue on RTIPPR (1)* (2010), page(s): 21–29.

[Hoa+17] Vinh Truong Hoang et al. « LBP Histogram Selection based on Sparse Representation for Color Texture Classification. » In : *VISIGRAPP (4 : VISAPP)*. 2017, page(s): 476–483.

[Hon+03] Anxiang Hong et al. « Region-of-interest based flower images retrieval ». In : *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*. T. 3. IEEE. 2003, page(s): III–589.

[HLC11] Tzu-Hsiang Hsu, Chang-Hsing Lee et Ling-Hwei Chen. « An interactive flower image recognition system ». In : *Multimedia Tools and Applications* **53**.*1* (2011), page(s): 53–73.

[Hu62] Ming-Kuei Hu. « Visual pattern recognition by moment invariants ». In : *IRE transactions on information theory* **8**.*2* (1962), page(s): 179–187.

[Hua+06a] Zhi-Kai Huang et al. « Bark classification based on gabor filter features using rbpnn neural network ». In : *Neural Information Processing*. Springer. 2006, page(s): 80–87.

[Hua+06b] Zhi-Kai Huang et al. « Bark classification based on textural features using artificial neural networks ». In : *Advances in Neural Networks*. Springer, 2006, page(s): 355–360.

[ILP13] Mohamed Issolah, Diane Lingrand et Frédéric Precioso. « SIFT, BoW Architecture and one-against-all Support Vector Machine. » In : *CLEF (Working Notes)*. 2013.

[Jid+16] Lv Jidong et al. « Recognition of apple fruit in natural environment ». In : *Optik-International Journal for Light and Electron Optics* **127**.*3* (2016), page(s): 1354–1362.

[Jol+16a] Alexis Joly et al. « Implicit biodiversity monitoring from mobile search logs ». In : *Proceedings of the ACM Multimedia Conference 2016*. CONFERENCE-2017-166. 15-19 October 2016. 2016.

[Jol+16b] Alexis Joly et al. « LifeCLEF 2016 : multimedia life species identification challenges ». In : *International Conference of the Cross-Language Evaluation Forum for European Languages*. Springer. 2016, page(s): 286–310.

[Jol+17] Alexis Joly et al. « Lifeclef 2017 lab overview : multimedia species identification challenges ». In : *International Conference of the Cross-Language Evaluation Forum for European Languages*. Springer. 2017, page(s): 255–274.

[KWT88] Michael Kass, Andrew Witkin et Demetri Terzopoulos. « Snakes : Active contour models ». In : *International journal of computer vision* **1**.*4* (1988), page(s): 321–331.

[Kaz+14] Wajahat Kazmi et al. « Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors : Analysis and comparison ». In : *ISPRS journal of photogrammetry and remote sensing* **88** (2014), page(s): 128–146.

[KYU10] Hanife Kebapci, Berrin Yanikoglu et Gozde Unal. « Plant image retrieval using color, shape and texture features ». In : *The Computer Journal* **54**.*9* (2010), page(s): 1475–1490.

[Kha+16] Haider Adnan Khan et al. « Abnormal mass classification in breast mammography using rotation invariant LBP ». In : *Electrical Engineering and Information Communication Technology (ICEEICT), 2016 3rd International Conference on*. IEEE. 2016, page(s): 1–5.

[Koh+11] Armin Kohan et al. « Robotic harvesting of rosa damascena using stereoscopic machine vision ». In : *World Applied Sciences Journal* **12**.*2* (2011), page(s): 231–237.

[KSH12] Alex Krizhevsky, Ilya Sutskever et Geoffrey E Hinton. « Imagenet classification with deep convolutional neural networks ». In : *Advances in neural information processing systems*. 2012, page(s): 1097–1105.

[LP13] Thi-Lan Le et Ngoc-Hai Pham. « MICA at ImageClef 2013 Plant Identification Task. » In : *CLEF (Working Notes)*. 2013.

[Le+15] Thi-Lan Le et al. « Complex background leaf-based plant identification method based on interactive segmentation and kernel descriptor ». In : *Proceedings of the 2nd International Workshop on Environmental Multimedia Retrieval*. ACM. 2015, page(s): 3–8.

[LH17] Hyo-Haeng Lee et Kwang-Seok Hong. « Automatic recognition of flower species in the natural environment ». In : *Image and Vision Computing* **61** (2017), page(s): 98–114.

[LS91] Bing-Cheng Li et Jun Shen. « Fast computation of moment invariants ». In : *Pattern Recognition* **24**.*8* (1991), page(s): 807–813.

[LLH11] Peilin Li, Sang-heon Lee et Hung-Yao Hsu. « Review on fruit harvesting method for potential use of automatic fruit harvesting systems ». In : *Procedia Engineering* **23** (2011), page(s): 351–366.

[Man+01] A-G Manh et al. « AE—Automation and emerging technologies : Weed leaf image segmentation by deformable templates ». In : *Journal of agricultural engineering research* **80**.*2* (2001), page(s): 139–146.

[MYA16] Mostafa Mehdipour Ghazi, Berrin Yanikoğlu et Erchan Aptoula. « Open-set plant identification using an ensemble of deep convolutional neural networks ». In : CLEF. 2016.

[Mil09] Julien Mille. « Narrow band region-based active contours and surfaces for 2D and 3D segmentation ». In : *Computer Vision and Image Understanding* **113**.*9* (2009), page(s): 946–965.

[MA04] Farzin Mokhtarian et Sadegh Abbasi. « Matching shapes with self-intersections : Application to leaf classification ». In : *IEEE Transactions on Image Processing* **13**.*5* (2004), page(s): 653–661.

[MB95] Eric N Mortensen et William A Barrett. « Intelligent scissors for image composition ». In : *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM. 1995, page(s): 191–198.

[ML09] Marius Muja et David G Lowe. « Fast approximate nearest neighbors with automatic algorithm configuration. » In : *VISAPP (1)* **2**.*331-340* (2009), page(s): 2.

[Mzo+16] Olfa Mzoughi et al. « Semantic-based automatic structuring of leaf images for advanced plant species identification ». In : *Multimedia Tools and Applications* **75**.*3* (2016), page(s): 1615–1646.

[NZ12] Asma Najjar et Ezzeddine Zagrouba. « Flower image segmentation based on color analysis and a supervised evaluation ». In : *Communications and Information Technology (ICCIT), 2012 International Conference on*. IEEE. 2012, page(s): 397–401.

[Nak13] Hideki Nakayama. « NLab-UTokyo at ImageCLEF 2013 Plant Identification Task. » In : *CLEF (Working Notes)*. Citeseer. 2013.

[NLB12] Loris Nanni, Alessandra Lumini et Sheryl Brahnam. « Survey on LBP based texture descriptors for image classification ». In : *Expert Systems with Applications* **39**.*3* (2012), page(s): 3634–3641.

[Neg+17] Tamiris Trevisan Negri et al. « A Robust Descriptor for Color Texture Classification Under Varying Illumination. » In : *VISIGRAPP (4 : VISAPP)*. 2017, page(s): 378–388.

[Ngu+17] Thi Thanh-Nhan Nguyen et al. « A combination of deep learning and hand-designed feature for plant identification based on leaf and flower images ». In : *Asian Conference on Intelligent Information and Database Systems*. Springer. 2017, page(s): 223–233.

[NZ06] Maria-Elena Nilsback et Andrew Zisserman. « A visual vocabulary for flower classification ». In : *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. T. 2. IEEE. 2006, page(s): 1447–1454.

[NZ07] Maria-Elena Nilsback et Andrew Zisserman. « Delving into the Whorl of Flower Segmentation. » In : *BMVC*. 2007, page(s): 1–10.

[NZ08] Maria-Elena Nilsback et Andrew Zisserman. « Automated flower classification over a large number of classes ». In : *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*. IEEE. 2008, page(s): 722–729.

[OPH96] Timo Ojala, Matti Pietikäinen et David Harwood. « A comparative study of texture measures with classification based on featured distributions ». In : *Pattern recognition* **29**.*1* (1996), page(s): 51–59.

[OPM02] Timo Ojala, Matti Pietikainen et Topi Maenpaa. « Multiresolution gray-scale and rotation invariant texture classification with local binary patterns ». In : *IEEE Transactions on pattern analysis and machine intelligence* **24**.*7* (2002), page(s): 971–987.

[Ots79] Nobuyuki Otsu. « A threshold selection method from gray-level histograms ». In : *IEEE transactions on systems, man, and cybernetics* **9**.*1* (1979), page(s): 62–66.

[Pac+14] Dénes Paczolay et al. « Wlab of university of szeged at lifeclef 2014 plant identification task ». In : *Working notes of CLEF 2014 conference*. 2014.

[Pal04] Christoph Palm. « Color texture classification by integrative co-occurrence matrices ». In : *Pattern recognition* **37**.*5* (2004), page(s): 965–976.

[PJJ12] HN Patel, RK Jain et MV Joshi. « Automatic segmentation and yield measurement of fruit using shape analysis ». In : *International Journal of Computer Applications* **45**.*7* (2012), page(s): 19–24.

[PG01] Alessio Plebe et Giorgio Grasso. « Localization of spherical fruits for robotic harvesting ». In : *Machine Vision and Applications* **13**.*2* (2001), page(s): 70–79.

[PVM07] Alice Porebski, Nicolas Vandenbroucke et Ludovic Macaire. « Iterative feature selection for color texture classification ». In : *Image Processing, 2007. ICIP 2007. IEEE International Conference on*. T. 3. IEEE. 2007, page(s): III–509.

[Por+14] Alice Porebski et al. « A new benchmark image test suite for evaluating colour texture classification schemes ». In : *Multimedia Tools and Applications* **70**.*1* (2014), page(s): 543–556.

[RKB04] Carsten Rother, Vladimir Kolmogorov et Andrew Blake. « Grabcut : Interactive foreground extraction using iterated graph cuts ». In : *ACM transactions on graphics (TOG)*. T. 23. 3. ACM. 2004, page(s): 309–314.

[Rou13] Georges Rouy. *Flore de France, ou Description des plantes qui croissent spontanément en France, en Corse et en Alsace-Lorraine*. Société des sciences naturelles de la Charente-Inférieure, 1893-1913.

[Rus+15] Olga Russakovsky et al. « Imagenet large scale visual recognition challenge ». In : *International Journal of Computer Vision* **115**.*3* (2015), page(s): 211–252.

[SAK04] Takeshi Saitoh, K Aoki et Toyohisa Kaneko. « Automatic Recognition of Blooming Flowers ». In : *Proceedings - International Conference on Pattern Recognition*. T. 1. Sept. 2004, page(s): 27–30 Vol.1. ISBN : 0-7695-2128-2.

[SK00] Takeshi Saitoh et Toyohisa Kaneko. « Automatic recognition of wild flowers ». In : *Pattern Recognition, 2000. Proceedings. 15th International Conference on*. T. 2. IEEE. 2000, page(s): 507–510.

[SK03] Takeshi Saitoh et Toyohisa Kaneko. « Automatic recognition of wild flowers ». In : *Systems and Computers in Japan* **34**.*10* (2003), page(s): 90–101.

[Sch16] John H Schaffner. « A general system of floral diagrams ». In : *Ohio Journal of Science* (1916).

[Sel+17] Ramprasaath R Selvaraju et al. « Grad-CAM : Visual Explanations from Deep Networks via Gradient-Based Localization. » In : *ICCV*. 2017, page(s): 618–626.

[Sob14] Irwin Sobel. « An Isotropic 3x3 Image Gradient Operator ». In : (fév. 2014).

[Suz+85] Satoshi Suzuki et al. « Topological structural analysis of digitized binary images by border following ». In : *Computer vision, graphics, and image processing* **30**.*1* (1985), page(s): 32–46.

[Šva14] M Švab. « Computer-vision-based tree trunk recognition ». 2014. URL : `http://www.vicos.si/Downloads/TRUNK1p`.

[Sze+15] Christian Szegedy et al. « Going deeper with convolutions ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, page(s): 1–9.

[TZF02] Teruo Takahashi, Shuhuai Zhang et Hiroshi Fukuchi. « Measurement of 3-D locations of fruit by binocular stereo vision for apple harvesting in an orchard ». In : *2002 ASAE Annual Meeting*. American Society of Agricultural et Biological Engineers. 2002, page(s): 1.

[TKC09] Chin-Hung Teng, Yi-Ting Kuo et Yung-Sheng Chen. « Leaf segmentation, its 3d position estimation and leaf classification from a few images with very close viewpoints ». In : *International Conference Image Analysis and Recognition*. Springer. 2009, page(s): 937–946.

[Vap99] Vladimir Naumovich Vapnik. « An overview of statistical learning theory ». In : *IEEE transactions on neural networks* **10**.*5* (1999), page(s): 988–999.

[VRK11] Sara Vicente, Carsten Rother et Vladimir Kolmogorov. « Object cosegmentation ». In : *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE. 2011, page(s): 2217–2224.

[Wan+04] Yuan-Yuan Wan et al. « Bark texture feature extraction based on statistical texture analysis ». In : *Intelligent Multimedia, Video and Speech Processing, 2004. Proceedings of 2004 International Symposium on*. IEEE. 2004, page(s): 482–485.

[Wan+00] Zhiyong Wang et al. « Leaf image retrieval with shape features ». In : *International Conference on Advances in Visual Information Systems*. Springer. 2000, page(s): 477–487.

[WSG11] Andreas Wendel, Sabine Sternig et Martin Godec. « Automated identification of tree species from images of the bark, leaves and needles ». In : *16th Computer Vision Winter Workshop*. Citeseer. 2011, page(s): 67.

[Wil17] K.J. Willis. *State of the World's Plants 2017*. Rapp. tech. Royal Botanic Gardens, Kew, 2017.

[Woj11] Michael Wojtech. *Bark : A Field Guide to Trees of the Northeast*. University Press of New England, 2011.

[Wu+07] Stephen Gang Wu et al. « A leaf recognition algorithm for plant classification using probabilistic neural network ». In : *Signal Processing and Information Technology, 2007 IEEE International Symposium on*. IEEE. 2007, page(s): 11–16.

[Xia+15] Chunlei Xia et al. « In situ 3D segmentation of individual plant leaves using a RGB-D camera for agricultural automation ». In : *Sensors* **15**.*8* (2015), page(s): 20463–20479.

[XJY14] Rong Xiang, Huanyu Jiang et Yibin Ying. « Recognition of clustered tomatoes based on binocular stereo vision ». In : *Computers and Electronics in Agriculture* **106** (2014), page(s): 75–90.

[Zha+14] Yudong Zhang et al. « Fruit classification using computer vision and feedforward neural network ». In : *Journal of Food Engineering* **143** (2014), page(s): 167–177.

[ZTK05] Jun Zhao, Joel Tow et Jayantha Katupitiya. « On-tree fruit recognition using texture properties and color data ». In : *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE. 2005, page(s): 263–268.

[ZN04] Jie Zou et George Nagy. « Evaluation of model-based interactive flower recognition ». In : *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. T. 2. IEEE. 2004, page(s): 311–314.

Table des figures

1 Illustrations de différentes interfaces de l'application Folia de reconnaissance d'espèces d'arbres via une photo de feuille sur smartphone. 4

1.1 Photo d'un chêne pédonculé et photo de plants de vulpins des prés (épis). . 11
1.2 Hiérarchie taxonomique du vivant dans la classification classique (CC BY-SA 3.0 via Wikimedia Commons) 13
1.3 Les espèces du genre *Quercus* donnent des glands (ici *Quercus cerris*, *Quercus coccifera*, *Quercus ilex*, *Quercus robur*). 13
1.4 Fleurs de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae* 14
1.5 Fruits de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae* 14
1.6 Extrait de la flore de Bonnier [Bon08]. 15
1.7 Extrait de la flore de Coste [Cos03]. 16
1.8 Évolution de l'écorce de *Betula pendula* entre un arbre jeune (à gauche) et un arbre plus âgé (à droite) 16
1.9 Similitude entre l'écorce de *Quercus robur* à gauche et de *Acer platanoides* à droite, et entre une feuille de *Cornus mas* à gauche et de *Viburnum tinus* à droite 17
1.10 Schéma d'une fleur par Hervé Sauquet. 17
1.11 Fleur femelle et fleurs mâles du châtaignier (*Castanea sativa*). 18
1.12 Schéma des différents types d'inflorescences. 18
1.13 Illustration des deux types de symétrie chez les fleurs de *Prunus avium* (actinomorphe) et de *Cercis siliquastrum* (zygomorphe). 19
1.14 Fleur de pomme de terre (*Solanum tuberosum*) et diagramme floral dont la formule florale est K : (5) ; [C : (5) ; A : 5] ; Gsup : (2) 19
1.15 Fleurs de deux individus de *Nerium oleander* de cultivars différents. 20
1.16 Structure schématique d'un fruit typique. 21
1.17 Développement du fruit (gland) chez *Quercus pubescens*. 21
1.18 Variabilité intra-classe des feuilles au sein du mûrier à papier (*Broussonetia papyrifera*). 22

1.19 Schéma général d'une feuille d'Angiosperme. 22
1.20 Différentes familles de texture d'écorce selon Michael Wojtech [Woj11] avec mise en évidence des structures (dessin). 28
1.21 Exemple de lenticelles de différentes formes : rondes, allongées, losangiques. 28
1.22 Les fruits de de *Broussonetia papyrifera* a) et d'*Euonymus europaeus* b) sont confondus avec des fleurs à cause de leur aspect très coloré. Les chatons de *Carpinus betulus* c) et les bourgeons de fleurs de *Viburnum tinus* d) sont confondus avec des fruits. 29
1.23 Fleur (à gauche) et fruit (à droite) de *Platanus* x *hispanica*. Fleur (à gauche) et fruit (à droite) de *Betula pendula*. 29
1.24 Exemple de question pour le formulaire de distinction entre les fruits et les fleurs. 30
1.25 Résultats pour l'auto-évaluation du niveau botanique de l'utilisateur. . . . 31
1.26 Exemple de fleurs et de fruits bien reconnus par les utilisateurs. 31
1.27 Option fausse préférentiellement choisie entre le fruit et la fleur. 31
1.28 Option correcte préférentiellement choisie, mais grande incertitude. 32
1.29 Extrait du livre "Guide de la faune et de la flore" [Eis17]. 34
1.30 Interface des outils *ID-Botanica* (gauche) et *Key to Nature* (droite). 34
1.31 Illustration du logiciel IDAO 35
1.32 Captures d'écran des applications pour smartphone *Quelle est cette fleur ?* a), *Sauvages de ma rue* b), *IK-Botanic*c) et *Clés de forêt* d). 36

2.1 Captures d'écran de l'outil de reconnaissance de plantes de l'application *GFinger*. 40
2.2 Captures d'écran de l'application *Pl@ntNet* lors des différentes étapes d'identification d'une plante. 41
2.3 Captures d'écran de l'application *LeafSnap UK* lors des différentes étapes d'identification d'un arbre. 42
2.4 Captures d'écran de l'application *Folia* présentant les interfaces d'interaction utilisateur. 42
2.5 Interface pour la segmentation d'une feuille, avec un marqueur interne et des marqueurs externes, les bassins versants générés, et la feuille segmentée. Extrait de Le *et al.* [Le+15] 44
2.6 Boîte englobante (cadre extérieur) autour de la fleur à segmenter et détection, puis correction, des points de départ de la création du contour, extrait de Hsu *et al.* [HLC11] 45
2.7 Utilisation de l'algorithme des ciseaux intelligents [MB95] pour la segmentation d'une feuille, avec ajout de trois ancres lorsque le contour extrait ne suit pas bien le bord de la feuille. 45

2.8 Les différentes étapes de la segmentation d'une feuille avec GrabCut. Une première segmentation est calculée à partir de la ROI tracée par l'utilisateur (rectangle vert). L'utilisateur peut ensuite rajouter des parties manquantes (points rouges) et indiquer des parties n'appartenant pas à l'objet d'intérêt (points bleus). 46
2.9 Illustration de l'ajustement des paramètres de la courbe *rhodonea* pour la segmentation d'une fleur, issue de [ZN04]. 46
2.10 Les différents organes de *Acer campestre* sont parfois difficiles à isoler de l'arrière-plan. 48
2.11 Illustration des résultats de l'algorithme de segmentation 3D issue de [TKC09] avec une des images sources, les points 3D appairés et les feuilles segmentées. 49
2.12 Application d'un modèle paramétrique déformable de feuille pour la segmentation, issue de [Cer13b] 49
2.13 Résumé de l'algorithme de segmentation de Nilsback et Zisserman, issu de [NZ07] 50
2.14 Illustration de l'utilisation de l'aspect circulaire des fruits pour la segmentation, issue de [PJJ12] 51
2.15 Analyse de la forme de la fleur selon Saitoh *et al.*. Illustration issue de [SK03]. 53
2.16 Détection des dents sur le contour de la feuille avec deux valeurs de seuil différentes pour les maxima de courbure, issue de [Aro+12] 53
2.17 Illustration de la discrétisation de l'espace colorimétrique RGB avec 4 bins par canal, issue de [Zha+14]. 54
2.18 Vocabulaire de couleurs appris pour la reconnaissance de fruits ou de fleurs, issue de [Cer+13a]. 55
2.19 Visualisation des points SIFT sur différents organes végétaux. 56
2.20 Exemple d'arbre de décision. 57
2.21 Les fleurs (rangée du haut) ne ressemblent pas toujours à des fleurs, et les fruits (rangée du bas) ne sont pas toujours d'aspect circulaire et coloré. . . 61

3.1 Exemples d'écorces de *Quercus ilex* à différents âges, d'abord lisses puis crevassées. 68
3.2 Similarité inter-classes entre les arbres *Acer platanoide*, *Populus nigra*, *Quercus nigra*, et *Tilia platyphyllios*. 68
3.3 Les différentes familles de structures d'écorce basées sur des critères visuels, avec des dessins mettant en évidence les structures distinctives des écorces. 69
3.4 Exemple de recadrage d'une image d'écorce (image initiale à gauche et image recadrée à droite). 69
3.5 Visualisation du nombre d'images de la base pour chaque espèce. 70

3.6 Exemple de filtres de Gabor avec deux orientations θ et différentes valeurs de fréquences f. 71
3.7 Résultats des filtrages de l'image initiale a) avec une orientation de 0° b) et 90° c). 72
3.8 Résultat de l'algorithme de Canny pour extraire les contours horizontaux et verticaux des images d'écorces types de chaque famille. 74
3.9 Formation du "mot vertical" W_V. 74
3.10 Exemple de mots verticaux et horizontaux pour deux types d'écorce avant normalisation. 75
3.11 Représentation de l'espace colorimétrique HSV. 76
3.12 Exemples d'images d'écorces et leur histogramme associé. 77
3.13 Résultats de classification pour une base de 101 écorces d'arbres. 78
3.14 Évolution du taux de classification en fonction du nombre de neurones sur la couche cachée pour un réseau de neurones composé de trois couches. . . 79
3.15 Évolution du taux de classification en fonction du nombre de neurones sur les deux couches cachées pour un réseau de neurones composé de quatre couches. 80
3.16 Évolution du taux de classification des écorces selon la prise en compte ou non du vecteur caractéristique de Gabor. 81
3.17 Évolution du taux de classification en fonction du nombre de lignes et de colonnes du quadrillage pour la formation du mot vertical et du mot horizontal. 83
3.18 Histogramme sommé des canaux H pour les écorces de la base de données d'apprentissage. 84
3.19 Évolution du taux de classification global (Top 1) des images d'écorces classifiées uniquement avec le vecteur caractéristique de couleur en fonction de la taille de ce vecteur. 84
3.20 Évolution du taux de classification moyen par espèce des images d'écorces classifiées uniquement avec le vecteur caractéristique de couleur en fonction de la taille de ce vecteur. 85
3.21 Évolution des taux de classification globaux des images d'écorces classifiées avec le vecteur caractéristique complet en fonction de la taille de ce vecteur. 86
3.22 Évolution du taux de classification moyen par espèce des images d'écorces classifiées avec le vecteur caractéristique complet en fonction de la taille de ce vecteur. 86
3.23 Résultats de classification pour une base de 101 écorces d'arbres avec un histogramme de teintes réduit à 30 bins. 87

3.24 Exemple d'images de la base *NewBarkTex*. a) *Betula pendula* b) *Fagus silvatica*, c) *Picea abies*, d) *Pinus silvestris*, e) *Quercus robus* et f) *Robinia pseudacacia* 87

3.25 Taux de classification sur la base de données NewBarkTex avec le vecteur complet pour deux valeurs de quantification de l'histogramme H. 88

3.26 Taux de classification sur la base de 101 classes d'écorces avec différents descripteurs de textures. 89

3.27 Architecture du CNN AlexNet, illustration issue de [KSH12]. 90

3.28 Architecture du CNN GoogLeNet, illustration issue de [Sze+15]. 91

4.1 Variabilité de forme, de couleur et de texture des fruits (les trois images de gauche) et des fleurs (les trois images de droite) des arbres présents en France métropolitaine. 95

4.2 Co-segmentation du fruit de *Gleditsia triacanthos*. La première ligne présente les images en entrée de l'algorithme de Dong *et al.* [Don+15] avec en vert les marques sur le fond et en rouge les marques sur l'objet. La seconde ligne expose les résultats de segmentation de l'objet pour chaque image d'entrée. C'est la cinquième image qui présente un arrière-plan différent. . . 96

4.3 Schéma des deux caméras en face arrière de l'iPhone 7 Plus. 97

4.4 Captures d'écran de l'application permettant de prendre en photo un organe avec chacune des deux caméras arrières de l'iPhone 7 Plus, ainsi que l'interface permettant de dessiner la trace utilisateur sur l'objet d'intérêt. . 98

4.5 Exemples de photos prises avec a) la caméra "grand-angle", b) la caméra "téléobjectif", c) la trace utilisateur et d) l'image "téléobjectif" segmentée manuellement. 99

4.6 Image téléobjectif initiale a) avec la trace utilisateur b) et le masque de fond c). 100

4.7 Exemples de patchs conservés (verts) ou rejetés (rouges) pour la classe objet. 100

4.8 Carte de texture de l'image initiale a) avec un patch pour $n = 11$ en niveau de gris b) et en couleur c). 102

4.9 Autre exemples de carte de texture de l'image initiale a) avec un patch pour $n = 11$ en niveau de gris b) et en couleur c). 102

4.10 Comparaison des cartes de texture entre des patchs en niveau de gris avec $n = 11$ a) et $n = 5$ b), et des patchs couleurs avec $n = 11$ c) et $n = 5$ d). . 103

4.11 Exemples d'images initiales (a, c, e) avec le contour de la trace utilisateur en bleu, et les cartes de similarité de couleurs associées (b, d, f). 104

4.12 Amélioration des cartes de MLP grâce à la prise en compte de la carte de similarité de couleurs. Image initiale avec la trace utilisateur en bleu a), carte de similarité de couleurs b), carte de MLP avant suppression des patchs du fond ayant une couleur trop similaire à l'objet d'intérêt c), et carte de MLP après suppression des patchs d). 105
4.13 Exemples de cartes de MLP obtenues à partir des images "téléobjectif" et des traces utilisateur (contour bleu). 105
4.14 Schéma de la génération de la carte de MLP. 106
4.15 Illustration de CC_{MLP} avec l'image "téléobjectif" en a), l'image binaire en sortie du MLP $D_{MLP\,binary}$ b), les composantes connexes c), la composante connexe intersectant la trace utilisateur d) et la segmentation manuelle e). 106
4.16 Visualisation des indices c) pour comparer une région segmentée Ω_{seg} b) par rapport à la vérité terrain Ω_{obj} a). Illustration issue de [Cer13a] 107
4.17 Résultats de segmentation d) différents en fonction de la trace utilisateur en entrée a) avec la carte de MLP b) et CC_{MLP} c). 109
4.18 Influence de la dilatation de la trace utilisateur sur la méthode CC_{MLP}. . . 109
4.19 Influence de l'érosion de la trace utilisateur sur la méthode CC_{MLP}. 109
4.20 Extraction des points d'intérêts SURF sur l'image grand-angle en haut et l'image téléobjectif en bas a) sur l'ensemble de l'image b) et restreint au centre c). 111
4.21 Mise en correspondance des 30 paires de points caractéristiques présentant les meilleurs appariements. 112
4.22 Visualisation c) de l'application sur l'image grand-angle de la matrice d'homographie calculée entre l'image grand-angle a) et l'image téléobjectif b). 113
4.23 Flux optique entre l'image téléobjectif et l'image grand-angle transformée dans l'espace géométrique de l'image téléobjectif. 114
4.24 Image grand-angle b) et téléobjectif c) d'un petit fruit, suscitant une détection et un appariement des points d'intérêt en partie sur le fond a). La matrice d'homographie étant incorrecte, la transformation géométrique de l'image grand-angle ne centre pas l'objet d'intérêt d), donnant donc une carte de profondeurs e) non pertinente. 115
4.25 Appariement des points caractéristiques restreints sur la trace utilisateur. 116
4.26 Le calcul de la transformation géométrique entre les images a) et b) permet de générer l'image c) où l'objet d'intérêt est centré, permettant de calculer l'image de profondeur d) grâce à la norme du flux optique calculé entre a) et c). 116
4.27 Schéma de la génération de la carte de profondeurs D_P. 117

4.28 Influence de la surface de la trace utilisateur sur la carte de profondeurs sur la moyenne des pixels de la carte de profondeurs à l'intérieur de la trace utilisateur. 118
4.29 Influence de la surface de la trace utilisateur sur la carte de profondeurs sur l'écart-type des pixels de la carte de profondeurs à l'intérieur de la trace utilisateur. 118
4.30 Paire d'images du gland du chêne chevelu (*Quercus cerris*) photographié sous deux angles légèrement différents. 120
4.31 Cartes de profondeur obtenues pour la paire d'images de la Figure 4.30. . . 120
4.32 Fleur de grenadier (*Punica granatum*) photographiée avec différentes focales. 121
4.33 Résultats de cartes de profondeur pour les images de la Figure 4.32. 122
4.34 Carte de profondeurs entre les images 4.32a et 4.32b. 122
4.35 Évolution du modèle de contour actif initialisé sur la trace utilisateur a). L'image est en niveau de gris pour une meilleure visualisation. 125
4.36 Contribution de la carte de profondeurs par rapport à la carte de MLP seule. a) Image initiale, b) Carte de profondeurs D_P, c) Carte de MLP D_{MLP}, d) Segmentation contour actif avec D_{MLP} seule, e) Segmentation contour actif avec D_{MLP} et D_P. 127
4.37 Variation des indices de mesures de qualité de segmentation en fonction du poids ϵ de la carte de profondeurs D_P. 127

5.1 Schéma d'une figue. 134
5.2 Correspondances entre les différentes classes des organes et les espèces, genres et familles associés. 136
5.3 Les différents traitements avant l'extraction des moments de Hu avec l'image initiale a), l'image seuillée b) et les contours extraits c). 137
5.4 Deux inflorescences de même couleur pour lesquelles l'information de texture pourrait les différencier. 140
5.5 Calcul du LBP pour le pixel central (en rouge). 140
5.6 Répartition des images dans les différentes classes. 142
5.7 Image initiale (à gauche) et deux images augmentées par transformation de perspective. 143
5.8 Image initiale (à gauche) et image transformée (à droite) avec la visualisation du bruit ajouté. 144
5.9 Histogramme "sommé" des canaux H pour les fruits et les fleurs segmentés de la base de données d'apprentissage non augmentée. 146
5.10 Évolution du taux de classification global des fruits et des fleurs classifiés uniquement avec l'histogramme de teintes H en fonction de la taille de celui-ci. 147

5.11 Évolution du taux de classification par classe des fruits et des fleurs classifiés uniquement avec l'histogramme de teintes H en fonction de la taille de celui-ci. 147
5.12 Exemple de confusion entre les fruits et les fleurs de l'espèce *Betula pendula*. 149
5.13 Exemple de confusion entre les fruits des espèces *Acer pseudoplatanus* et *Acer saccharinum*. 149
5.14 Matrice de confusion issue de la prédiction des espèces. 151
5.15 Détail de la matrice de confusion sur la prédiction des espèces (Figure 5.14) pour les espèces appartenant au genre *Quercus*. 151
5.16 Matrice de confusion issue de la prédiction des genres. 152
5.17 Détail de la matrice de confusion sur la prédiction des genres (Figure 5.16) pour les genres appartenant à la famille des *Betulaceae*. 152
5.18 Fleurs de genres différents *Alnus*, *Betula*, *Carpinus* et *Corylus* appartenant à la même famille *Betulaceae*. 153
5.19 Matrice de confusion issue de la prédiction des familles. 153
5.20 Évolution du taux de classification pour la classification des classes de fruits-fleurs a), et pour la prédiction sur les espèces b), les genres c), les familles d). 154
5.21 Exemple d'images originales a), b), c) d), et leur image retaillée associée e), f), g), h) pour mettre en entrée des CNN. 155

6.1 Les différentes étapes de traitement pour l'identification de la feuille par l'application *Folia*. 161
6.2 Taux de reconnaissance des vecteurs $\mathbf{F}_l$, $\mathbf{F}_b$ et $\mathbf{F}_{l,b}$ avec le classifieur de *Folia*. 163
6.3 Taux de reconnaissance des vecteurs $\mathbf{F}_l$, $\mathbf{F}_b$ et $\mathbf{F}_{l,b}$ avec le classifieur SVM un-contre-tous. 163
6.4 Combinaison *A posteriori* du vecteur de feuille $\mathbf{F}_l$ et du vecteur d'écorce $\mathbf{F}_b$. 165
6.5 Matrice de confusion issue de la classification des feuilles grâce au classifieur de *Folia*. 167
6.6 Matrice de confusion issue de la classification des écorces grâce au classifieur SVM un contre tous. 168
6.7 Variabilité de la marge des feuilles du houx commun (*Ilex aquifolium* [Cer13b]). 169
6.8 Matrice de confusion issue de la classification avec le classifieur SVM des vecteurs de feuille et d'écorce concaténés. 170
6.9 Matrice de confusion issue de la combinaison *a posteriori* des vecteurs de feuille et d'écorce. 171
6.10 Exemples d'écorces de *Betula pendulas* commençant à un jeune âge et devenant de plus en plus âgées. 171
6.11 Exemples d'images d'écorces de *Quercus Ilex* présentant une grande variété de textures. 172

6.12 Confusion entre les feuilles de *Quercus ilex* a) et c) et les feuilles de *Laurus nobilis* b) et *Ilex aquifolium* d). 172
6.13 Résultats de classification globaux sur les écorces avant et après avoir pris en compte la matrice. 173
6.14 Résultats de classification moyens par espèce sur les écorces avant et après avoir pris en compte la matrice de confusion. 174
6.15 Résultats de classification globaux et moyens par espèce sur les feuilles avant et après avoir pris en compte la matrice de confusion. 174
6.16 Résultats de classification globaux sur les vecteurs de feuilles et d'écorces concaténés avant et après avoir pris en compte la matrice. 175
6.17 Résultats de classification moyens par espèce sur les vecteurs de feuilles et d'écorces concaténés avant et après avoir pris en compte la matrice de confusion. 175
6.18 Résultats de classification globaux sur les vecteurs de feuilles et d'écorces combinés *a posteriori* avant et après avoir pris en compte la matrice. . . . 176
6.19 Résultats de classification moyens par espèce sur les vecteurs de feuilles et d'écorces combinés *a posteriori* avant et après avoir pris en compte la matrice de confusion. 176
6.20 Résultats de classification globaux et moyens par espèce sur les fruits-fleurs avant et après avoir pris en compte la matrice de confusion. 178
6.21 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion. 179
6.22 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des écorces et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion. 179
6.23 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles, des écorces et des fruits-fleurs avant et après avoir pris en compte la matrice de confusion. 180
6.24 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres. 181
6.25 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des écorces et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres. 182
6.26 Résultats de classification globaux et moyens par espèce de la combinaison *a posteriori* des feuilles, des écorces et des fruits-fleurs avec la prédiction sur les espèces ou sur les genres. 182

7.1 Illustration par Frédéric PHAN des couches de convolutions du CNN AlexNet. 188
7.2 Exemples d'images de la partie "Leaf (scan and scan-like)" de la base de données PlantCLEF. 189
7.3 Images de la partie "Leaf (scan and scan-like)" de la base de données PlantCLEF avec du bruit gaussien sur les bords de manière à conserver le *ratio* de la forme des feuilles. 190
7.4 Échelle de couleur "Jet". 191
7.5 Sélection de feuilles pour visualiser les différentes couches de convolutions. 192
7.6 Visualisation des images en sortie des noyaux de convolution des cinq couches de convolution du CNN AlexNet sur une image de feuille (*Acer monspessulanum*). 193
7.7 Classification des filtres actifs et non actifs pour la première couche de convolution. 194
7.8 Filtres de la deuxième couche de convolutions séparant les contours des feuilles selon certaines orientations. 194
7.9 Visualisation du même filtre de la couche de convolution 3 pour différentes feuilles avec des formes de base différentes. 195
7.10 Visualisation de filtres caractéristiques de la quatrième couche de convolution du réseau AlexNet. 195
7.11 Visualisation du même filtre de la couche de convolution 5 pour différentes feuilles lobées ou non. 196
7.12 Feuille de *Acer monspessulanum* utilisée pour la visualisation des couches de convolutions. 197
7.13 Visualisation de la sortie des filtres de convolution de la couche numéro 3 du réseau AlexNet. 198
7.14 Détection des filtres de convolution actifs grâce à la méthode de binarisation. 199
7.15 Résultat de la détection de filtres actifs par calcul de la valeur moyenne des pixels. 200
7.16 Résultat de la détection de filtres actifs par calcul de l'écart-type de l'histogramme. 200
7.17 Résultat de détection des filtres actifs avec la méthode basée sur le seuillage d'Otsu. 202
7.18 Exemple d'espèces avec une caractéristique de base tronquée. 203
7.19 Niveau d'activation des filtres de la troisième couche de convolution pour des feuilles présentant une base tronquée. 204
7.20 Autres exemples de niveaux d'activation pour la troisième couche de convolution. 204
7.21 Activation des filtres de la troisième couche de convolution pour la base tronquée après soustraction de la moyenne. 205

7.22 Cartes de plausibilité obtenues à partir des présences théoriques des espèces (0 en blanc et 1 en vert foncé). Illustrations issues de [Cer13b]. 212

Liste des tableaux

1.1 Description des différents fruits et de leurs propriétés morphologiques. . . . 21
1.2 Les différentes formes générales de feuille. Issue de [Cer13b]. 23
1.3 Les différentes formes de feuille simple. Issue de [Cer13b]. 23
1.4 Les différentes formes de nervures. Issue de [Cer13b]. 24
1.5 Les différentes formes de bases de feuille. Issue de [Cer13b]. 24
1.6 Les différentes formes d'apex de feuille. Issue de [Cer13b]. 25
1.7 Les différentes formes de bord de feuille. Issue de [Cer13b]. 25
1.8 Les différents de feuilles composées. Issue de [Cer13b]. 26
1.9 Les différentes positions d'attache des feuilles à la tige. 26

3.1 Taux de classification sur la base d'écorces *NewBarkTex* avec différents vecteurs de caractéristiques. 88
3.2 Comparaison des résultats de classification sur la base de 101 classes d'écorces avec des CNN. 91

4.1 Mesure de qualité de la segmentation avec la méthode CC_{MLP}. 108
4.2 Mesures de qualité de segmentation pour les méthodes CC_{MLP} et AC_{MLP} prenant en compte uniquement la carte de MLP calculée sur les images dites téléobjectifs. 125
4.3 Comparaison des résultats de CC_{MLP} et (AC_{MLP} avec GAC pour la segmentation de feuilles. 126
4.4 Mesure de qualité de segmentation pour la segmentation via contour actif avec des poids δ et ϵ adaptés aux images. 128

5.1 Résultats de classification par SVM des fruits et des fleurs avec et sans augmentation de données. 144
5.2 Résultats de classification par MLP des fruits et des fleurs avec et sans augmentation de données. 145
5.3 Résultats de classification par MLP des fruits et des fleurs avec prédiction sur les espèces, les genres et les familles. 148
5.4 Performances des CNN AlexNet et GoogLeNet sur des images non segmentées. 155

5.5 Performances des CNN AlexNet et GoogLeNet sur des images segmentées avec, et sans, augmentation de données. 156

6.1 Évolution des taux de classification pour chaque espèce avant et après raffinement pour le vecteur d'écorce complet. 174
6.2 Résumé des taux de classification pour les différentes méthodes de combinaison des vecteurs de feuilles et d'écorces. 177
6.3 Résumé des taux de classification pour les différentes méthodes de combinaison des vecteurs de feuilles, d'écorces et de fruits-fleurs. 180
6.4 Résumé des taux de classification pour les combinaisons *a posteriori* des vecteurs de feuilles, d'écorces et de fruits-fleurs en prenant en compte la prédiction sur l'espèce ou sur le genre. 182