

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Batna 2

Faculté des Mathématiques et de l'Informatique
Département d'Informatique



THÈSE

pour obtenir le diplôme de
Doctorat en Informatique

**Identification et Classification des Espèces de Plante
à Partir des Images d'Écorce**

présentée et soutenue publiquement par

Safia BOUDRA

le 19/09/2021

Jury:

Président:	Pr. Kamel Eddine MELKEMI	Université de Batna 2
Rapporteur:	Pr. Ali BEHLOUL	Université de Batna 2
Co-Rapporteur:	Dr. Itheri YAHIAOUI	Université de Reims
Examineur:	Pr. Mohamed Chaouki BABAHENINI	Université de Biskra
Examineur:	Pr. Rachid SEGHIR	Université de Batna 2
Examineur:	Dr. Larbi GUEZOULI	Université de Batna 2

Remerciements

J'EXPRIME mes vifs remerciements aux membres de Jury : tout d'abord, M. Kamel Eddine MELKEMI, pour l'honneur qu'il m'a fait en acceptant de présider le jury, ensuite, M. Mohamed Chaouki BABAHENINI et M. Rachid SEGHIR et M. Larbi GUEZOULI pour avoir accepté la tâche laborieuse de rapporter ma thèse. Je les remercie tous pour l'intérêt et le temps qu'ils ont accordé à mon travail, et pour les différentes suggestions avisées pour l'améliorer.

Je remercie également mon encadreur M. Ali BEHLOUL pour avoir accepté d'encadrer ma thèse, pour toute son aide, sa disponibilité et pour ses conseils pertinents et son soutien tout au long de de cette thèse.

Je profite de cet espace pour faire part de mon immense gratitude envers mon co-encadrante Mme Itheri YAHIAOUI pour m'avoir accueilli au sein du laboratoire CReSTIC à l'université de Reims.

Je la remercie infiniment pour sa totale implication et accompagnement tout au long des travaux de ma thèse. Grâce à ses conseils précieux et avisés, sa rigueur scientifique, j'ai beaucoup appris durant ces années de thèse. Ces conseils m'ont guidé dans la réflexion, la critique et la rédaction scientifique. Elle était constamment et généreusement disponible tant sur le plan professionnel que moral malgré ses engagements. Elle a su répondre à toutes mes questions. Je la remercie pour ces innumérables relectures, et les discussions que nous avons menées durant toutes ces années. Sa patience, son écoute et ses inlassables encouragements ont profondément contribué dans mon parcours scientifique.

Je la remercie aussi pour m'avoir donné l'occasion de participer à plusieurs événements scientifiques autour de la thématique de la vision par ordinateur et la reconnaissance d'objets, pour m'avoir accordé l'accès aux supports nécessaires pour la réalisation de mes travaux durant mon séjour à Reims.

Enfin je remercie mes parents et ma famille pour leurs chaleureux encouragements durant les années de la préparation de ma thèse. Vous avez été toujours là pour m'épauler dans toutes les circonstances.

Résumé

L'IDENTIFICATION des espèces végétales est une étape clé pour la préservation de la biodiversité. Cette identification est essentiellement réalisée par des experts botanistes. Ils déploient un savoir en taxonomie pour la description des caractéristiques visuels, comme la forme, la couleur et la structure des différents organes de la plante. L'espèce sous-jacente est généralement déterminée via un processus de recherche par similarité avec les espèces connues. Cependant, cette approche trouve ces limites dans le très grand nombre d'espèces existantes, et dans les problèmes de similarité inter-espèces et de variabilité intra-espèce, qui sont assez présents pour ces objets naturels. Ceci favorise le développement d'outils d'aide à l'identification automatique des espèces de plante. Cette solution est réalisable grâce à la croisée des domaines de la vision par ordinateur et la botanique.

Cette thèse aborde le problème d'identification et classification automatique des espèces végétales à partir des images d'écorce. Le problème est formalisé comme un problème d'analyse de texture, et simule le processus de description et d'identification réalisé par les botanistes.

Nous proposons de décrire la texture d'écorce via l'extraction et la représentation des macrostructures qui la caractérisent. À l'aide d'une représentation multi-échelles combinée à une représentation de la distribution d'intensité, et différents encodages des variations inter-échelles de la distribution, nous dérivons des codes macro-motifs statistiques binaires caractéristiques pour les différentes structures d'écorce. Les images d'écorce sont représentées par leurs descripteurs multi-échelles de macro-motifs binaires. Différents classifieurs supervisés sont utilisés pour l'identification et la classification de leurs espèces. De plus, nous procédons à la classification automatique des espèces d'écorce par des réseaux de neurones à convolution. En contraste à notre première proposition, avec un système classique de reconnaissance d'image, il s'agit ici d'une approche basée sur l'apprentissage profond des caractéristiques des écorces.

Des évaluations réalisées sur quatre bases d'écorce montrent que notre approche permet d'identifier différentes espèces, dont les images présentent plusieurs problèmes photométriques ainsi que des problèmes de similarité inter-espèces. Nous avons aussi montré que notre approche est plus performante que les réseaux de neurones à convolution entièrement entraînés sur des petites bases d'écorce.

Mots clés : *Identification des plantes, écorce, texture, représentation multi-échelles, représentation statistique, macrostructure, macro-motifs binaires, réseaux de neurones à convolution.*

Table des matières

Table des matières	vii
Liste des figures	ix
Liste des tableaux	xiii
1 Introduction générale	1
1.1 Vers l'identification automatique des espèces végétales	2
1.2 Regard botanique et formalisation des caractéristiques visuelles	6
1.3 Objectifs	9
1.4 Plan de la thèse	11
2 État de l'art	13
2.1 Analyse de la texture des écorces	14
2.2 Classification par apprentissage profond	38
2.3 Synthèse et discussion	41
2.4 Contexte du travail	43
2.5 Conclusion	47
3 Modèle générique pour la représentation de la texture	49
3.1 Modèle de description structurale	51
3.2 Extraction de la macrostructure	52
3.3 Représentation de la macrostructure	55
3.4 Encodage de la macrostructure	57
3.5 Descripteur de la texture d'écorce	59
3.6 Conclusion	62
4 Descripteurs des macro-motifs binaires statistiques	63
4.1 Conception de la macrostructure	64
4.2 Distribution d'intensité	69
4.3 Codes macro-motifs	76
4.4 Descripteurs de la texture d'écorce	81
4.5 Propriétés des codes macro-motifs	82
4.6 Conclusion	84

5 Bases d'écorce et protocole d'évaluation	85
5.1 Bases d'écorce	86
5.2 Classification et mesures d'évaluation	89
5.3 Méthodes en comparaison	91
5.4 Conclusion	92
6 Résultats expérimentaux	93
6.1 Étude comparative des variantes LBP multi-échelles	94
6.2 Évaluation des descripteurs de macro-motifs binaires statistiques	100
6.3 Classification des écorces par les CNN	133
6.4 Conclusion	138
7 Conclusions et Perspectives	139
7.1 Résumé des travaux	140
7.2 Perspectives	141

Liste des figures

1.1	Les problèmes de déformation morphologique, et la complexité du fond naturel pour l'identification des espèces de feuilles, fleurs, et fruits.	4
1.2	Avantages de l'identification à partir de l'écorce : surface rigide, facile à isoler du fond naturel, et riche en information structurale.	4
1.3	Challenges et problèmes rencontrés lors de l'analyse de l'apparence de l'écorce pour l'identification de son espèce.	5
1.4	Structure d'écorce (images (a) et (b) WOJTECH [2011]).	7
1.5	Les types botaniques de l'écorce. Les types sont définis selon l'apparence de la structure présente sur la surface de l'écorce WOJTECH [2011].	8
1.6	Exemples de structures caractérisantes de certains types d'écorce WOJTECH [2011].	8
2.1	Exemple de calcul de la matrice de cooccurrences. (a) Image en entrée. (b) Matrice de cooccurrences pour un déplacement $(d, \theta) = (1, 0^\circ)$ TUCERYAN et JAIN [1993].	16
2.2	Les étapes de génération du descripteur LBP.	18
2.3	Décomposition en ondelettes (2DDWT).	20
2.4	Décomposition en ondelettes d'une image d'écorce en 2 niveaux.	20
2.5	Exemple de calcul du vecteur d'intersections BERTRAND et collab. [2017].	22
2.6	Détecteur et descripteur de points d'intérêt SIFT LOWE [2004].	23
2.7	Espace d'échelle de MBLBP LIAO et collab. [2007], et construction du descripteur multi-échelles.	27
2.8	La variante multi-échelles LBP (image (a) prise de MÄENPÄÄ et PIETIKÄINEN [2003]).	27
2.9	La variante multi-échelles MSLBP (image (a) prise de MÄENPÄÄ et PIETIKÄINEN [2003]).	27
2.10	La variante multi-échelles PLBP (images prises de QIAN et collab. [2011]).	27
2.11	Les filtres utilisés pour générer les image réponses pour les construction des dictionnaire des textons VARMA et ZISSERMAN [2005].	30
2.12	Contraintes qui limitent l'application à l'identification et la classification d'écorce.	46
3.1	Exemple de représentation d'une structure de texture d'écorce. Un large support spatial permet de représenter l'information de la macrostructure, tandis qu'un support très local ne le permet pas.	50
3.2	Combinaison d'applications pour l'extraction d'une macrostructure de \mathcal{L} niveaux.	53
3.3	Distribution des niveaux de l'espace d'échelle.	54
3.4	Distribution des points voisins.	54

3.5	Exemple de différentes macrostructures avec $\mathcal{L} = 3$ niveaux. (a) Distribution dense pour les niveaux d'échelle et des points voisins. (b) Distribution dense pour les niveaux d'échelle et creuse pour des points voisins. (c) Distribution grossière pour les niveaux d'échelle et dense des points voisins. (d) Distribution grossière pour les niveaux d'échelle et creuse des points voisins.	55
3.6	Représentation des niveaux d'échelle de la macrostructure.	56
3.7	Combinaison d'applications pour l'encodage de la macrostructure et construction des <i>macro-motifs</i>	58
3.8	Représentation de la texture selon différents niveaux de descriptions : global, par région, et local par des points d'intérêt.	60
4.1	Une macrostructure circulaire de trois niveaux.	65
4.2	Distribution des niveaux d'échelle selon $\tau = \{1, 2, 3\}$	65
4.3	Localité du voisinage avec différentes valeurs de $R_1 R_1 \in [1, 6]$, et $\tau = 1$	65
4.4	Une macrostructure de trois anneaux avec deux cercles chacun.	66
4.5	Largeur d'anneau.	66
4.6	Macrostructure circulaire et macrostructure en anneaux avec différentes largeurs. Exemple de comparaison en terme de taille d'espace d'échelle et support spatial pour les deux représentations de l'espace d'échelle.	67
4.7	Distribution dense des points voisins <i>AllN</i>	68
4.8	Distribution creuse des points voisins <i>OneN</i>	69
4.9	Exemple de description statistique avec $\mathcal{L} = 3$	76
4.10	Calcul des codes macro-motifs radiaux $SRBP_m$	77
4.11	Calcul des prototypes statistiques intrinsèques.	79
4.12	Calcul des codes macro-motifs profonds $SMBP_m$	81
4.13	Calcul des histogrammes des descripteurs $SRBP_m$ et $SMBP_m$. Exemples avec une macrostructure de $\mathcal{L} = 3$ niveaux d'échelle.	82
5.1	Exemples d'écorce de chaque classe des bases d'écorce.	88
6.1	Courbes de Rappel/Précision des variantes LBP multi-échelles sur les bases Trunk12 et AFE.	96
6.2	Visualisation des images LBPF des réponses retournées à une image d'écorce requête.	98
6.3	Taux d'identification (%) et valeurs MAP des descripteurs singuliers $SMBP_m$ et leur combinaison SMBP avec les prototypes statistiques μ^{scale} et $\mu^{pattern}$ sur la base Trunk12. La configuration utilisée est $\{R_1 = 1, \tau = 1, AllN\}$	101
6.4	Différence entre les valeurs des MAP par espèce pour le descripteur SMBP avec μ^{scale} et celle du descripteur SMBP avec $\mu^{pattern}$	102
6.5	Variabilité intra-espèce pour l'espèce #3 de la base Trunk12.	102
6.6	Image requête de l'espace #3 et les trois premières images retournées avec leurs représentations statistiques pour les différents prototypes.	102
6.7	Comparaison des scores d'identification (%) et des valeurs de MAP pour les descripteurs SRBP et SMBP avec les configurations <i>circ</i> et <i>ring</i> sur la base Trunk12.	107

6.8 Valeurs MAP par espèce des meilleures configurations des différents descripteurs statistiques, avec <i>AfIN</i> sur la base Trunk12.	108
6.9 Taux d'erreur de confusion de l'espèce #4 avec les espèces #1, #9, #11, #12 de la base Trunk12.	108
6.10 Taux d'identification (%) et valeurs MAP des combinaisons formées de deux descripteurs statistiques sur la base Trunk12.	110
6.11 Comparaison des scores d'identification (%) et des valeurs de MAP pour les descripteurs SRBP et SMBP avec les configurations <i>circ</i> et <i>ring</i> sur la base AFF.	115
6.12 Taux d'identification (%) et valeurs MAP des combinaisons formées de deux descripteurs statistiques sur la base AFF.	116
6.13 Matrices de confusion des descripteurs circ-SMBP et MRELBP [LIU et collab., 2016b] sur la base AFF.	119
6.14 Exemple de la même image requête de l'espèce #3 de la base AFF et ses réponses selon circ-SMBP et MRELBP.	119
6.15 Exemple d'écorce de l'espèce #11 de la base AFF dont la surface est couverte de lichen.	119
6.16 Comparaison des scores d'identification en 1-NN et SVM _{rbf} pour les descripteurs SRBP et SMBP avec les configurations <i>circ</i> et <i>ring</i> sur la base Bark101.	123
6.17 Taux de classification (%) en 1-NN et SVM _{rbf} des combinaisons formées de deux descripteurs statistiques sur la base Bark101.	124
6.18 Comparaison des scores de classification (%) en 1-NN et SVM _{lin} pour les descripteurs SRBP et SMBP avec les configurations <i>circ</i> et <i>ring</i> sur la base BarkNet.	129
6.19 Taux de classification (%) en 1-NN et SVM _{lin} des combinaisons formées de deux descripteurs statistiques sur la base BarkNet.	130
6.20 Matrices de confusion des descripteurs circ-SMBP et MRELBP [LIU et collab., 2016b] avec SVM _{lin} sur la base BarkNet.	132

Liste des tableaux

2.1	Résumé des travaux sur les écorces.	15
2.2	Résumé des différentes variantes LBP.	35
5.1	Information générale sur les bases d'écorce.	88
6.1	Les valeurs des paramètres des différentes variantes LBP multi-échelles en fonction du niveau de l'échelle.	95
6.2	Valeurs de MAP pour les variantes LBP basique et MResLBP sur les bases Trunk12 et AFF	95
6.3	Valeurs de MAP pour les variantes LBP multi-échelles sur la base Trunk12.	96
6.4	Valeurs de MAP pour les variantes LBP multi-échelles sur la base AFF	96
6.5	Taux d'identification (%) du descripteur circ-SRBP sur la base Trunk12 avec différentes représentations de la macrostructure.	104
6.6	Taux d'identification (%) du descripteur circ-SMBP sur la base Trunk12 avec différentes représentations de la macrostructure.	105
6.7	Taux d'identification (%) des descripteurs ring-SRBP et ring-SMBP sur la base Trunk12 avec différentes représentations de la macrostructure.	107
6.8	Configurations retenues pour les différents descripteurs statistiques sur la base Trunk12.	108
6.9	Taux d'identification (%) et valeurs MAP des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base Trunk12.	110
6.10	Résultats de comparaison avec l'état de l'art sur la base Trunk12. les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].	111
6.11	Taux d'identification (%) du descripteur circ-SRBP sur la base AFF avec différentes représentations de la macrostructure.	113
6.12	Taux d'identification (%) du descripteur circ-SMBP sur la base AFF avec différentes représentations de la macrostructure.	114
6.13	Taux d'identification (%) des descripteurs ring-SRBP et ring-SMBP sur la base AFF avec différentes représentations de la macrostructure.	116
6.14	Configurations retenues pour les différents descripteurs statistiques sur la base AFF	116
6.15	Taux d'identification (%) et valeurs MAP des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base AFF	117
6.16	Résultats de comparaison avec l'état de l'art sur la base AFF les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].	118

6.17 Taux de classification (%) en 1-NN du descripteur circ-SRBP sur la base Bark101 avec différentes représentations de la macrostructure.	121
6.18 Taux de classification (%) en 1-NN du descripteur circ-SMBP sur la base Bark101 avec différentes représentations de la macrostructure.	122
6.19 Taux de classification (%) en 1-NN des descripteurs ring-SRBP et ring-SMBP sur la base Bark101 avec différentes représentations de la macrostructure.	123
6.20 Configurations retenues pour les différents descripteurs statistiques sur la base Bark101.	123
6.21 Taux de classification (%) en 1-NN et SVM _{rbf} des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base Bark101.	124
6.22 Résultats de comparaison avec l'état de l'art sur la base Bark101. Les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].	125
6.23 Taux de classification (%) en 1-NN du descripteur circ-SRBP sur la base BarkNet avec différentes représentations de la macrostructure.	127
6.24 Taux de classification (%) en 1-NN du descripteur circ-SMBP sur la base BarkNet avec différentes représentations de la macrostructure.	128
6.25 Taux de classification (%) en 1-NN des descripteurs ring-SRBP et ring-SMBP sur la base BarkNet avec différentes représentations de la macrostructure.	130
6.26 Configurations retenues pour les différents descripteurs statistiques sur la base BarkNet.	130
6.27 Taux de classification (%) en 1-NN et SVM _{lin} des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base BarkNet.	131
6.28 Résultats de comparaison avec l'état de l'art sur la base BarkNet.	131
6.29 Résultats de la classification des espèces d'écorce avec ResNet34 entièrement entraîné sur les bases d'écorce.	134
6.30 Taux de classification des images d'écorce avec ResNet34 pré-entraîné et réadapté sur les bases d'écorce.	134
6.31 Comparaison des taux de classification de nos descripteurs SMBP et SRBP et ceux des descripteurs extraits de ResNet34 entièrement entraîné sur les bases d'écorces.	136
6.32 Comparaison des taux de classification de nos descripteurs SMBP et SRBP et ceux des descripteurs extraits de ResNet34 pré-entraîné et réadapté sur les bases d'écorce.	137

Chapitre 1

Introduction générale

Sommaire

1.1 Vers l'identification automatique des espèces végétales	2
1.2 Regard botanique et formalisation des caractéristiques visuelles	6
1.2.1 Structure d'écorce	6
1.2.2 Types d'écorce	8
1.2.3 Caractéristiques visuelles	9
1.3 Objectifs	9
1.4 Plan de la thèse	11

1.1 Vers l'identification automatique des espèces végétales

LES conséquences de la vie moderne et de l'industrialisation pèsent lourdement sur l'écosystème et vont de pair avec le changement climatique drastique pour affecter directement la biodiversité des espèces vivantes. Il est donc nécessaire de proposer des solutions efficaces et durables afin de protéger la biodiversité. Ces solutions passent par l'acquisition d'un savoir qui permettra de comprendre, d'identifier, et par la suite de préserver les espèces, notamment, les espèces végétales qui sont présentes dans la vie quotidienne de l'homme tant comme une source d'oxygène et d'aliments, que comme ressources dans divers secteurs de la vie comme l'agriculture, la pharmacie, le textile, etc.

On estime un nombre avoisinant les ~ 400.000 espèces végétales réparties sur plus de 450 familles avec en moyenne, presque 2000 espèces découvertes chaque année WILLIS [2017]. Ainsi, l'identification des espèces végétales constitue en elle seule un pilier important pour la préservation de la biodiversité.

L'identification est basée sur un savoir bien fondé en botanique et récemment, par l'étude du génome de la plante. Elle est élaborée par des experts taxonomistes et botanistes. Cela reste insuffisant et la tâche d'identification devient de plus en plus fastidieuse et difficile à réaliser, car basée sur un savoir centralisé en botanique et en taxonomie, avec des techniques bien complexes.

Une solution pour faciliter et accélérer l'identification des espèces végétales, est de mettre en disposition des botanistes des collections de spécimens, et des outils avancés qui permettent l'accessibilité à des informations botaniques. Durant les décennies passées, des efforts colossaux déployés à l'échelle mondiale ont permis d'émerger des volumes importants de données botaniques, et cela grâce à une campagne massive de digitalisation des collections des muséums, comme les travaux de [BELHUMEUR et collab., 2008] qui ont permis de numériser la collection complète du *Type Specimen Collection of the US National Herbarium* avec 90.000 spécimens¹. [BELHUMEUR et collab., 2008] ont par la suite développé un outil numérique pour l'aide à l'identification des espèces de plantes. Cet outil est actuellement utilisé par le *Smithsonian Institution National Museum on Natural History*.

Dans la même perspective, on cite également les initiatives de digitalisation du *New York Botanical Garden*² avec 4 millions d'images de spécimens et le *Missouri Botanical Garden*³ avec plus de 50.000 images de plantes en USA, le *Royal Botanical Gardens, Kew*⁴ en Grande Bretagne avec 50.000 images, et le *Muséum National d'Histoire Naturelle*⁵ en France avec presque 5 million d'images [LE BRAS et collab., 2017], et aussi *The Encyclopedia of Life*⁶ dont le but est de cataloguer, dans les années à venir, toutes les espèces d'organismes présentes sur terre.

En conséquence de la digitalisation des spécimens, et l'accessibilité aux bases d'images, il devient possible de formaliser le savoir en botanique par des outils d'aide à la reconnaissance automatique. Ces outils sont capables d'extraire de l'information à partir des images de plantes, et de

1. <https://naturalhistory.si.edu/research/botany>

2. <http://sweetgum.nybg.org/science/digital-collections/>

3. <http://www.worldfloraonline.org/>

4. <http://brahmsonline.kew.org/UKOT/Explore>

5. <https://science.mnhn.fr/institution/mnhn/collection/p/item/search>

6. <http://www.eol.org>

l'utiliser pour l'aide à la prise de décision quant à leurs espèces végétales. Cela devient de plus en plus réalisable grâce à l'avancée technologique dans les domaines de vision par ordinateur et de classification d'images.

Le développement des systèmes automatiques pour l'aide à l'identification des espèces végétales détient une importance particulière dans la communauté de vision par ordinateur. De nombreux systèmes, sous forme d'application web ou mobile ont émergé, on cite l'application mobile LeafSnap⁷ KUMAR et collab. [2012] qui est dédiée à l'identification des plantes à partir des images de feuilles explicitement disposées sur un fond blanc uniforme. Cette application couvre les 185 espèces de plantes de la flore Nord-Est américaine. L'application Flora Finder⁸ qui couvre les espèces natives de la Nouvelle-Zélande. L'application Folia⁹ CERUTTI et collab. [2013] dédiée à l'identification de plus d'une centaine d'espèces d'arbres et d'arbustes de la flore française à partir de l'analyse de la feuille sans imposer un fond uniforme. Aussi Pl@ntNet¹⁰ GOËAU et collab. [2011]; MOUINE et collab. [2013b] pour une identification multi-organes de plusieurs centaines d'espèces. L'application MobileFlora¹¹ propose une aide à l'identification des espèces à partir des images de fleurs. L'outil de recherche de la flore électronique eFlore¹² du réseau des botanistes français TelaBotanica¹³, qui permet de rechercher des informations précises sur une espèce végétale à partir de son nom scientifique ou son nom courant.

Tous ses systèmes fournissent un moyen de partage et d'approfondissement des connaissances botaniques, aussi bien pour les experts botanistes que les amateurs de la nature, comme les jardiniers et les forestiers. Ces derniers, ne disposent pas forcément d'un savoir préalable en botanique.

Toutefois, il est primordial de définir des restrictions qui cadrent le contexte d'utilisation de ces outils pour permettre un usage efficace. En effet, il ne peut y exister un outil universel capable d'analyser toutes les caractéristiques botaniques d'une espèce donnée, ou qui puisse identifier toutes les espèces existantes.

Ces spécifications portent sur l'organe utilisé (feuille, fleur ou écorce) à partir duquel on peut décider de l'espèce végétale de la plante. L'identification des plantes est généralement basée sur l'analyse des caractéristiques visuelles de ses organes. La grande majorité des travaux se sont focalisés sur l'utilisation des feuilles en raison de l'abondance de cet organe BELHUMEUR et collab. [2008]; COPE et collab. [2012]; WÄLDCHEN et MÄDER [2018]. Cependant, le défi majeur qui revient le plus souvent lors de la réalisation d'un système automatique d'identification à partir des images de feuilles est la présence du fond naturel dans l'image, ce qui nécessite une étape de segmentation a priori, qu'elle soit automatique ou semi-automatique CERUTTI et collab. [2013]. De plus, cet organe peut subir des changements morphologiques très conséquents, dus au développement de la feuille mais aussi aux conditions naturelles, comme les déformations non rigides qui peuvent altérer sa forme, critère visuel basique pour l'identification de l'espèce d'une feuille (c.f Figure 1.1). Par ailleurs, la plupart des feuilles tombent et ne sont pas présentes toute l'année, ce qui limite leur

7. <http://leafsnap.com/>

8. <http://www.florafinder.com>

9. <http://liris.univ-lyon2.fr/reves/>

10. <http://identify.plantnet-project.org/>

11. <http://www.mobileflora.com/MobileFlora.html>

12. <https://www.tela-botanica.org/flore/>

13. <https://www.tela-botanica.org/>



FIGURE 1.1 – Les problèmes de déformation morphologique, et la complexité du fond naturel pour l'identification des espèces de feuilles, fleurs, et fruits.



FIGURE 1.2 – Avantages de l'identification à partir de l'écorce : surface rigide, facile à isoler du fond naturel, et riche en information structurale.

utilisation pour l'identification de l'espèce végétale en question. Par conséquent, d'autres organes sont utilisés, notamment les organes reproductifs comme les fleurs et les fruits. Cependant, les mêmes problèmes naturels persistent pour ces organes aussi.

Un intérêt particulier a été porté pour un organe, peu utilisé avant dans les systèmes automatiques d'identification végétales, qu'est l'écorce [BERTRAND et collab. \[2017\]](#); [CARPENTIER et collab. \[2018\]](#); [CHI et collab. \[2003\]](#); [ŠULC et MATAS \[2013, 2017\]](#); [WENDEL et collab. \[2011\]](#). Cela revient aux propriétés de cet organe qui le rendent pratiquement convenable pour cette application. En effet, l'écorce est le seul organe présent tout au long de la durée de vie de la plante, aussi bien quand elle perd tous ses autres organes. De plus, l'écorce a une surface rigide qui permet une acquisition facile d'une image 2D. Tout cela, en plus de son accessibilité, et du fait que l'acquisition d'une image de petite taille soit riche en informations structurales suffisantes pour déterminer son espèce. Ceci peut ne pas nécessiter une étape de segmentation pour isoler un fond naturel potentiellement complexe (c.f Figure 1.2).

Cependant, étant un objet naturel, l'écorce est sujette à différents problèmes, ce qui rend l'application d'identification des espèces d'écorce relativement délicate et difficile à cerner. On cite dans la suite les différents problèmes *(i)* naturels, *(ii)* photométriques qui surviennent suite à la difficulté de contrôler les conditions d'acquisition d'image, et *(iii)* les problèmes de forte similarité inter-espèces et forte variabilité intra-espèce (c.f Figure 1.3) :

- ▶ Les problèmes naturels et environnementaux affectent particulièrement la surface de l'arbre et par conséquent l'apparence de l'écorce. On rencontre le plus souvent les traces de blessures, où une quantité importante de l'information structurale de l'écorce est perdue et ne

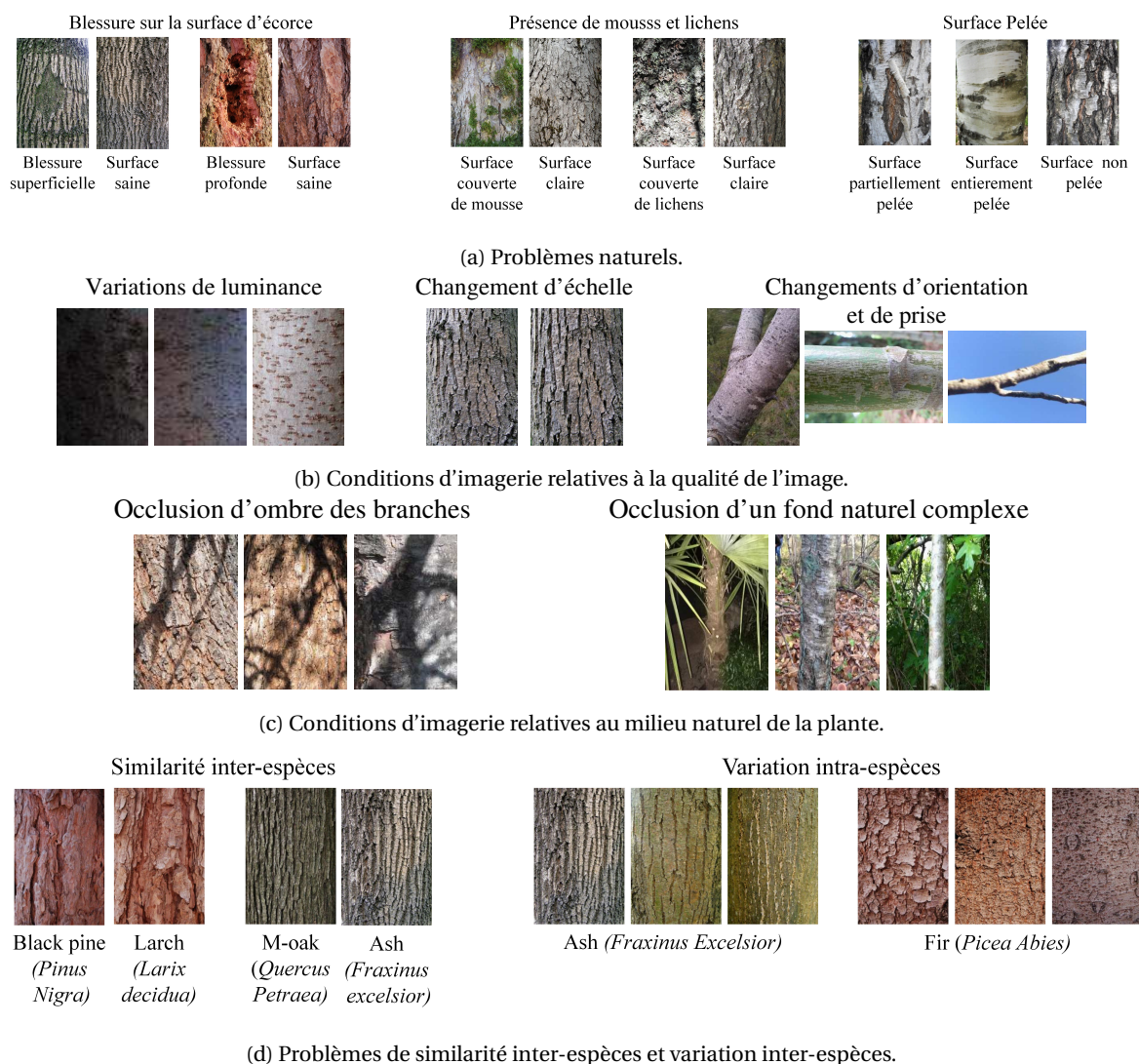


FIGURE 1.3 – Challenges et problèmes rencontrés lors de l'analyse de l'apparence de l'écorce pour l'identification de son espèce.

peut être restituée. Aussi, la présence d'organes vivants sur la surface du tronc de l'arbre comme la mousse et les lichens qui peuvent couvrir partiellement ou totalement la surface de l'écorce, et ainsi cacher la vraie apparence de l'écorce en introduisant une information bruitée. Pour certaines espèces, comme l'espèce *Betula*, on rencontre le problème des écorces pelées (c.f Figure 1.3a). En effet, avec une variation assez importante entre un échantillon pelé et un autre préservé, il est pratiquement difficile d'obtenir un seul modèle de représentation des deux surfaces.

- Les problèmes d'imagerie sont essentiellement causés par la difficulté de standardiser les conditions photométriques lors de la prise des images d'écorce dans leurs milieux naturels, comme pour le cas des applications effectuées sous des conditions de 'labo', car il s'agit là d'application sur des objets naturels. Les problèmes d'imagerie soulevés ici sont généralement relatifs aux changements d'échelle, i.e., la distance de prise de l'image par rapport au tronc de l'arbre, les changements de luminance, i.e., l'image est prise à différents moments de la journée et à différentes saisons (c.f Figure 1.3b). De plus, il y a éventuellement, le pro-

blème de changement d'orientation et de point de vue, et aussi une possible présence de l'ombre des autres organes, le plus souvent, des branches, ou d'un fond naturel complexe (c.f Figure 1.3c).

- Les problèmes relatifs à l'ambiguïté de similarité sont essentiellement présents dans les applications qui réalisent une identification de plusieurs classes d'un même objet (*fine grained applications*), comme le cas ici de l'objet écorce. Ainsi, et en raison de la vaste diversité des espèces vivantes existantes, une ambiguïté de similarité fait souvent face lors de l'automatisation du processus de description et décision des classes (ou bien espèces dans notre contexte).

La similarité inter-espèces se définit pour des espèces d'écorce qui partagent une apparence visuelle assez similaire malgré qu'elles appartiennent à des classes différentes. La variabilité intra-espèce se définit par la présence d'une variabilité de l'apparence de certains individus de la même espèce végétale indépendamment de l'âge de la plante ou de l'habitat naturel de ses échantillons. Cette ambiguïté peut être causée par des facteurs environnementaux. On montre dans la Figure 1.3d un exemple d'une écorce de l'espèce *Ash* (*Fraxinus excelsior*) qui est visiblement différente de certains individus de sa propre classe, mais partage une similarité d'apparence avec l'écorce de *M-oak* (*Quercu petraea*) qui est d'une espèce différente de la sienne. Cela rend difficile la différenciation de ces espèces, même à l'œil nu.

1.2 Regard botanique et formalisation des caractéristiques visuelles

1.2.1 Structure d'écorce

D'un point de vue botanique [GODET \[2012\]](#); [VAUCHER \[1993\]](#); [WOJTECH \[2011\]](#), l'écorce est défini comme un ensemble de couches superposées qui s'étalent de la surface jusqu'à une couche très fine interne appelée *le cambium vasculaire* comme le montre la Figure 1.4a à travers une coupe transversale d'un tronc d'arbre. Cet ensemble de couches à la particularité de se détacher du *bois* du tronc et des branches de l'arbre.

Le cambium vasculaire est une couche cylindrique de cellules qui se divisent et engendrent ainsi d'autres cellules. À la suite de cette division, une couche de *bois* se forme vers l'intérieur, et une couche de *phloème* (ou liber) se forme vers l'extérieur (c.f Figure 1.4b). Ainsi, ces cellules nouvellement produites par le cambium constituent le cycle d'accroissement en diamètre de l'arbre : la couche de bois se charge de la conduction ascendante de l'eau et des nutriments des racines vers les feuilles, et la couche de phloème actif transporte les produits générés par la photosynthèse (comme source de nourriture) dans toute la plante.

La partie supérieure du phloème actif constitue le *périderme* qui protège l'arbre des conditions environnementales comme la chaleur, le froid, et les insectes. À son tour, le périderme est constitué de trois couches; une couche intermédiaire, *le cambium cortical* qui se présente comme un ensemble de tissus cellulaires actifs qui produit *le phelloderme* vers l'intérieur, et *le liège* vers l'extérieur. Plusieurs couches fines de liège peuvent être produites par le cambium cortical durant l'année et servent de peau protectrice pour le tronc et les branches contre le dessèchement, les bactéries et les moisissures.

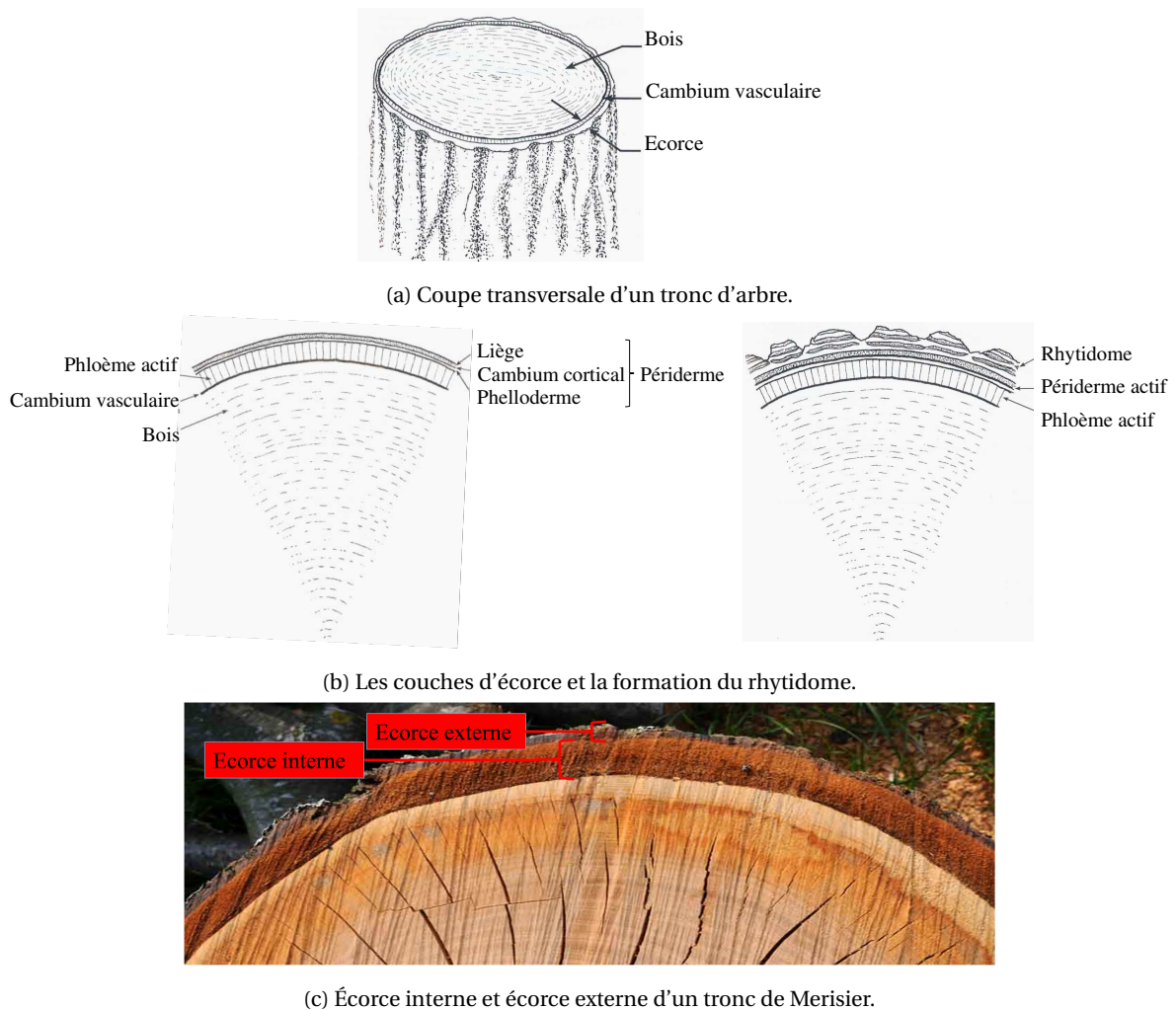


FIGURE 1.4 – Structure d'écorce (images (a) et (b) WOJTECH [2011]).

Plus l'arbre devient mature, plus la couche initiale du périoderme se dégénère et une nouvelle couche se forme depuis l'intérieur de l'ancien périoderme, dit *périoderme actif*. Cette couche se répartie sur tout le périmètre du tronc de l'arbre, de façon uniforme ou superposée. L'ensemble des cellules se trouvant vers l'extérieur du périoderme actif s'isole et meurt. À la succession de ce processus, une couche protectrice formée de l'ancien périoderme et des cellules mortes du phloème donne le *rhytidome* (c.f Figure 1.4b). C'est le rhytidome qui va donner à l'écorce son aspect définitif.

Ceci dit, l'écorce est donc constituée de deux parties : une partie vivante dite *écorce interne* (formée du phloème actif, phelloderme, et cambium cortical), et une partie morte dite *écorce externe* (formée de liège, périoderme, et rhytidome) comme le montre la Figure 1.4c.

L'épaisseur, la texture, et la forme de l'écorce externe dépendent de la réaction de chaque espèce à la pression exercée depuis l'intérieur lorsque de nouvelles couches de bois se forment. Pour certaines espèces, des fissures et craquement se forment, pour d'autre, l'écorce reste lisse.

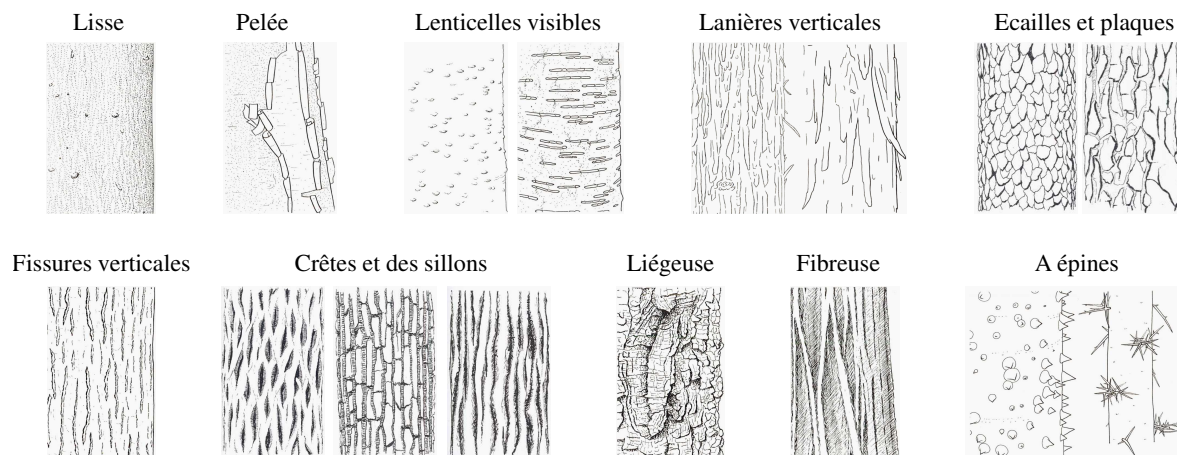


FIGURE 1.5 – Les types botaniques de l'écorce. Les types sont définis selon l'apparence de la structure présente sur la surface de l'écorce [WOJTECH \[2011\]](#).

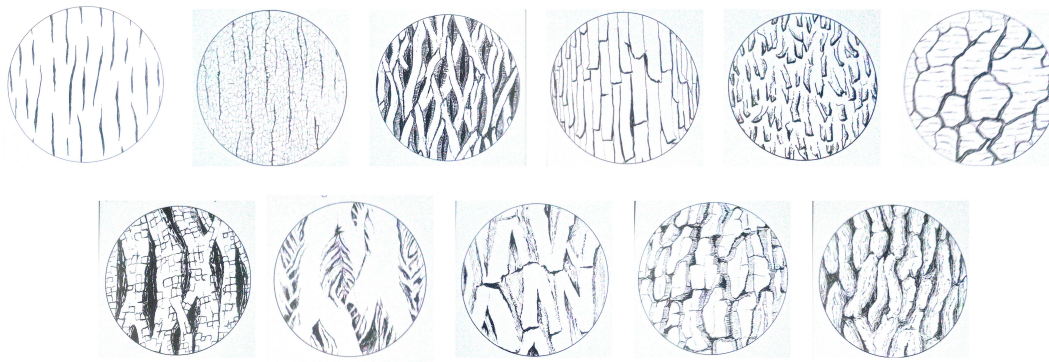


FIGURE 1.6 – Exemples de structures caractérisantes de certains types d'écorce [WOJTECH \[2011\]](#).

1.2.2 Types d'écorce

Selon la structure de l'écorce, on distingue dix types botaniques différents de l'écorce [VAUCHER \[1993\]](#); [WOJTECH \[2011\]](#) (c.f Figure 1.5) :

1. *Lisse (smooth)* : cette écorce est assez fine, dépourvue de fissures ou d'écailles.
2. *Pelée (peeling)* : le tronc est revêtu d'une écorce lisse est papyracée parcourue par des stries horizontales très apparentes.
3. À *Lenticelles visibles (lenticels visible)* : les lenticelles sont une sorte de 'pores' à travers lesquels l'arbre échange des gaz avec son environnement. Pour certaines espèces, ces lenticelles sont très visibles, et se présentent sous forme linéaire ou circulaire.
4. À *Lanières* ou plaquettes (*vertical strips*) : des stries verticales fines bien visibles se forment sur la surface de l'écorce et se détachent en lanières étroites.
5. *Écailles et plaques (scale and plates)* : écorce irrégulièrement craquelée, avec des gerçures en écailles de petite ou moyenne taille, ou bien avec des plaques de taille variable.
6. avec des *Fissures verticales (vertical cracks)* : présence de fissures ou nervures verticales peu profondes, des fois craquelées.
7. avec des *Crêtes et des sillons (ridges and furrows)* : présence de crêtes et sillons assez profonds. On retrouve généralement trois variantes dans ce type : des fissures entrelacées, des

fissures profondes en secteurs carrés ou rectangulaires, ou bien des sillons verticaux profonds continus à la crête bien formée.

8. *Liégeuse (corky)* : pour ce type, l'apparence de l'écorce est très irrégulière et le rhytidome est protubérant et dur.
9. *Fibreuse (fibrous)* : écorce spongieuse, à texture en crête longitudinale mais s'effritant en surface.
10. *À épines (thorn)* : ce type d'écorce est présent chez certaines espèces tropicales avec différentes formes d'épines sur la surface.

Ainsi, les botanistes et les naturalistes se basent essentiellement sur l'apparence de la structure d'écorce pour identifier son espèce ou sa famille. Une autre caractéristique reste informative mais pas pour autant discriminante, est la couleur de l'écorce. On retrouve généralement des nuances de gris-argenté, gris-vert, gris-noir, brun-gris, brun-rougeâtre, brun-verdâtre, blanc-grisâtre, blanc-jaunâtre, etc, qui peuvent varier de claires à foncées.

1.2.3 Caractéristiques visuelles

D'un point de vue traitement de contenu des images d'écorce, la plupart des approches proposées dans la littérature formalisent le problème d'analyse de la surface d'écorce par un problème *d'analyse de la primitive de la texture* CHI et collab. [2003]; ŠULC et MATAS [2017]; WAN et collab. [2004]; WENDEL et collab. [2011], combiné éventuellement avec la primitive de *la couleur* HUANG et collab. [2006c,d]. Cependant, la couleur n'est pas un attribut pertinent en soi, car très variable selon les conditions environnementales de la plante.

1.3 Objectifs

L'objectif d'un processus classique d'un système d'identification automatique de plantes est de pouvoir, à partir d'une photographie d'un ou plusieurs organes donnés, retourner son espèce végétale et éventuellement des informations complémentaires. La présente étude s'intéresse à l'identification automatique des espèces végétales à partir des images d'écorce. Deux principales phases constituent un tel système : une phase d'extraction d'information qui doivent être suffisamment discriminantes pour pouvoir aider à la prise de décision et retourner un sous ensemble d'espèces d'écorce auquel cette image est susceptible d'appartenir, ce qui constitue la seconde phase qu'est la classification.

Plus spécifiquement, on cherche par la phase d'extraction d'information à partir des images d'écorce à modéliser la description de la surface d'écorce. L'intérêt est de fournir une description spécifique des éléments caractéristiques des différentes espèces, à partir desquels, on peut les différencier. La Figure 1.6 montre de plus près ces structures caractérisantes de la texture pour quelques espèces d'écorce.

Ainsi, notre démarche se focalise essentiellement sur la partie *extraction et description* des éléments pertinents de la surface d'une écorce. On modélise ce problème par un problème d'analyse de texture, où on cherche à encoder une représentation de la structure, pour les différents types d'écorce, par une description de la *macrostructure* de la texture d'écorce. En effet, en se référant

aux sources de description d'une écorce pour l'identification de son espèce **GODET [2012]**; **VAUCHER [1993]**; **WOJTECH [2011]**, on s'aperçoit que la surface d'écorce est formée de structures avec une taille grande et variable, et un arrangement non-uniforme et non périodique (c.f Figure 1.6). Cependant, il est pratiquement difficile de définir un modèle mathématique qui modélise l'arrangement et la distribution des structures sur la surface d'écorce, car fortement affectés par les conditions naturelles et l'habitat de cette plante. De ce fait, on cherche plutôt à pouvoir définir un modèle qui simule l'apparence de ces structures en représentant et encodant l'information qu'elles englobent.

Ce modèle, sur lequel repose la principale contribution de cette étude, est inspiré de la description des structures de texture à base de micros motifs binaires (Local Binary Pattern (LBP)), où on définit une nouvelle variante multi-échelles capable d'encoder les différentes macrostructures présentes sur la surface d'écorce.

À l'encontre des variantes LBP multi-échelles, qui génèrent un code micro-motif par échelle spatiale, et ainsi un descripteur final de dimension très importante, avec le risque d'encoder des informations bruitées sur de larges supports spatiaux, l'approche définie dans cette étude a pour objectif de produire une représentation multi-échelles réduite de la macrostructure de l'écorce. Cette représentation vise à encoder l'information entre les échelles spatiales et générer un seul code macro-motif binaire représentatif de l'information de texture contenue dans la macrostructure. Des outils de composition d'applications, d'échantillonnage de l'information d'intensité dans l'espace d'échelle, de la statistique descriptive, et d'encodage binaire en profondeur permettent de formaliser la nouvelle représentation. Grâce à cela, cette représentation bénéficie d'une dimensionnalité réduite et des propriétés d'invariance au bruit et aux changements de rotation locale.

L'approche empruntée pour la prise de décision est une approche classique assez proche du processus effectué par les botanistes lors de l'identification d'une nouvelle espèce, i.e., retourner une liste d'espèces similaires avec lesquelles, l'espèce requête est censée partager des caractéristiques visuelles similaires et donc potentiellement y appartenir. Cette phase est donc supervisée et une vérité terrain permet de valider la décision de l'espèce recherchée, et par la suite de mesurer les performances de notre système.

Toutefois, ceci ne constitue pas une solution unique et universelle pour extraire tous les détails et aspects pertinents de la structure d'écorce, ou qui puisse couvrir la large panoplie des espèces de plantes existantes. Néanmoins, notre système d'identification offre un outil automatique d'aide à l'identification des espèces de plante à partir de leurs écorces dans un cadre de travail bien déterminé.

Autre que la conception 'classique' d'un système d'identification automatique des espèces de plante comme on vient de le présenter, on trouve des solutions entièrement basées sur l'apprentissage automatique des caractéristiques pertinentes, sans pour autant que ces dernières soient conçues explicitement par des méthodes classiques de vision par ordinateur. Il s'agit là de système à base de réseaux de neurones à convolution *Convolutional Neural Network (CNN)*, où la représentation du contenu visuel et la classification des différentes classes sont réalisées de bout en bout. L'utilisation des CNN et leurs comparaisons avec la solution proposée constituent la seconde contribution de cette étude. En effet, avec l'expansion des bases d'images, qui deviennent de plus en plus volumineuses, il convient d'élaborer une étude comparative quantitative afin de

pouvoir examiner les limites de la solution classique, et ainsi d'établir un ordre de grandeur entre les performances des deux approches.

1.4 Plan de la thèse

Le présent manuscrit est organisé comme suit :

- ▶ **Chapitre 2** : présente une revue de la littérature. Cette revue est composée de deux parties. La première partie est consacrée aux travaux réalisés sur la classification des espèces d'écorce. La seconde partie porte essentiellement sur une revue des travaux sur la classification de la texture générique. Dans cette partie, on expose les différentes approches de description de texture en mettant l'accent sur les approches structurales dérivées d'un encodage basé sur les motifs binaires à base d'une représentation locale ou une représentation multi-échelles. Ensuite, nous présentons une synthèse et une analyse des approches existantes afin de positionner et argumenter notre proposition pour une description de la texture des images d'écorce.
- ▶ **Chapitre 3** : dans ce chapitre, nous proposons un modèle générique de notre approche. La mise en œuvre de la modélisation proposée est détaillée grâce à l'utilisation du concepts de composition d'applications. Des exemples des représentations possibles de chaque phase du modèle générique sont exposées également. Ceci permet de dériver une famille de descripteurs différents selon les représentations utilisées.
- ▶ **Chapitre 4** : une représentation spécifique tirée à partir du modèle générique est présentée dans ce chapitre. L'idée principale est de générer une représentation multi-échelles réduite à base de macro-motifs binaires. Les détails de l'encodage de l'information de texture dans la macrostructure sont présentée, et quatre familles de descripteurs de texture sont introduits.
- ▶ **Chapitre 5** : dans cette partie, on présente l'ensemble des bases de données des images d'écorce utilisées pour l'évaluation de nos travaux. Ensuite, on présente le protocole d'identification ainsi que les mesures de performances pour interpréter et expliquer le comportement de nos descripteurs sur les bases d'écorce.
- ▶ **Chapitre 6** : ce chapitre est consacré à l'étude des performances des approches proposées et la validation de nos descripteurs de texture d'écorce dans le cadre d'un système d'identification des espèces végétales. Une étude comparative des différentes variantes multi-échelles LBP est exposée en premier lieu, elle sert à valider l'utilisation d'une approche structurale pour la description de la texture d'écorce. Ensuite, l'évaluation des familles de descripteurs proposés est effectuée sur différentes bases d'écorce suivie d'une comparaison avec un modèle CNN.
- ▶ **Chapitre 7** : ce chapitre présente la conclusion générale de nos travaux ainsi que les perspectives pour de futures travaux.

Chapitre 2

État de l'art

Sommaire

2.1 Analyse de la texture des écorces	14
2.1.1 Représentation de la texture d'écorce	16
2.1.2 Représentation de la couleur d'écorce	36
2.1.3 Classification des écorces	37
2.2 Classification par apprentissage profond	38
2.2.1 Réseaux de neurones à convolution CNN	39
2.2.2 Utilisation des CNN	40
2.2.3 Classification des écorces par CNN	41
2.3 Synthèse et discussion	41
2.4 Contexte du travail	43
2.5 Conclusion	47

L'IDENTIFICATION automatique des plantes consiste à analyser les caractéristiques visuelles des organes de la plante afin d'extraire des attributs informatifs pouvant mener à l'identification de son espèce végétale. Dans ce contexte, on retrouve des approches d'identification de plantes utilisant individuellement les feuilles [BELHUMEUR et collab. \[2008\]](#); [COPE et collab. \[2012\]](#); [MOUINE et collab. \[2013a\]](#); [MZOUGHU et collab. \[2016\]](#), les fleurs [NILSBACK et ZISSERMAN \[2010\]](#); [SEELAND et collab. \[2017\]](#) les fruits [AIBINU et collab. \[2011\]](#); [ZAWBAA et collab. \[2014\]](#), les aiguilles (pour les conifères) [WENDEL et collab. \[2011\]](#), les écorces [BERTRAND et collab. \[2017\]](#); [FEKRI-ERSHAD \[2020\]](#); [HUANG et collab. \[2006c\]](#); [REMEŠ et HAINDL \[2019\]](#); [ŠULC et MATAS \[2013, 2017\]](#); [WAN et collab. \[2004\]](#); [WENDEL et collab. \[2011\]](#), et ceux utilisant une combinaison de multiples organes [BERTRAND et collab. \[2017\]](#); [GOËAU et collab. \[2012, 2015, 2013, 2014a\]](#); [KIM et collab. \[2011\]](#); [ZHAO et collab. \[2015\]](#); [ZHU et collab. \[2017\]](#). Nous portons un intérêt spécifique à l'identification des espèces de plantes à partir des images d'écorce.

Ce chapitre est structuré en deux parties majeures pour l'analyse des images d'écorce. La première partie est dédiée aux approches de description du contenu visuel suivant le processus classique de représentation puis classification. La représentation génère un descripteur qui passe en entrée à des algorithmes d'apprentissage supervisé, i.e., la classification, pour déterminer la classe de l'image d'écorce. La seconde partie porte sur des approches d'apprentissage profond, où la classification des espèces est réalisée de bout-en-bout.

Ceci dit, nous détaillons en particulier la première partie, plus précisément, la phase de description du contenu visuel des images d'écorce, qui est essentiellement formulée comme un problème de représentation de texture. Nous présentons une revue des travaux sur la représentation des écorces. Ensuite, nous présentons un aperçu des approches génériques pour la représentation de la texture, suivie d'une synthèse et une critique qui expose les limites des approches actuelles. Ainsi, cela nous permettra d'élaborer les hypothèses retenues pour notre approche et aussi de la positionner parmi l'existant.

2.1 Analyse de la texture des écorces

En pratique, les botanistes identifient une espèce d'écorce par sa structure [VAUCHER \[1993\]](#); [WOJTECH \[2011\]](#). La structure est décrite en fonction de son arrangement, sa couleur, et la présence ou non d'autres éléments clés, comme les lenticelles, qui permettent d'approcher le type de cette écorce et par la suite d'identifier son espèce. Cependant, la tâche de description automatique de la structure de l'écorce n'est pas aussi évidente face à l'irrégularité, et la variabilité de la taille, et l'apparence de cette structure, principalement dues à des conditions naturelles. Or, il est toutefois possible de modéliser cette description à l'aide des méthodes de traitement d'image en les adaptant à encoder des attributs pertinents. Le problème de classification des écorces est essentiellement formulé comme un problème de classification de texture.

La texture est une caractéristique visuelle de la surface des objets. Elle reflète souvent un arrangement spatial d'éléments primitifs appelés *texons*, ou bien des microstructures fondamentales. Par ailleurs, une région de texture a des caractéristiques statistiques relatives à la périodicité, la répétition, et l'emplacement de ces textons, donnant ainsi différentes apparences à la surface des objets.

TABLEAU 2.1 – Résumé des travaux sur les écorces.

Représentation spatiale locale				Classification
Texture		Couleur		
WAN et collab. [2004]	GLCM	GLCM par plan	RGB	k-NN
LAMEI et collab. [2010]	GLCM	—	—	Genetic Programming (GP)
BRESSANE et collab. [2016]	GLCM	GLCM par plan	V de HSV	Random Forest (RF)
ZHAO et collab. [2015]	GLCM	—	—	SVM
BRESSANE et collab. [2015b]	GLCM+image statistique	GLCM + stats par plan	V de HSV	k-NN
HUANG et collab. [2006d]	GLCM+Dimension Fractal	Gabor par plan	YIQ	RBPNN
BLANCO et collab. [2016]	LBP	—	—	SVM
BAKIĆ et collab. [2013]	LBP ^{ri}	—	—	SVM
GOËAU et collab. [2014b]	LBP ^{ri}	—	—	k-NN
LE et collab. [2015]	noyau LBP	—	—	SVM
Représentation spatiale multi-échelles				Classification
Texture		Couleur		
CHI et collab. [2003]	Gabor	—	—	k-NN
HUANG et collab. [2006b]	Gabor	Distribution des angles CDA	RGB	RBPNN
BERTRAND et collab. [2017]	Canny+GWs+Gabor	Histogramme	H de HSV	SVM
HUANG [2006]	DWT	—	—	RBPNN
KIM et collab. [2011]	DWT	—	—	k-NN
SONG et collab. [2004]	DWT+LCLE+GLCM	—	—	k-NN
TSCHIEDEL GUEDES et BORGES [2018]	DT-CWT	—	—	k-NN
HUANG et collab. [2006c]	CT	CT par plan	YIQ	RBPNN
OTHMANI et collab. [2013]	DWT+DT-RCWF+CT	—	—	Random Forest (RF)
WENDEL et collab. [2011]	SIFT+BoW	—	—	SVM
ZHU et collab. [2017]	SIFT+BoW	—	—	SVM
GOËAU et collab. [2012]	Harris+SURF	Histogramme	HSV	k-NN-LOO
NAKAYAMA [2013]	SIFT (dense)	SIFT par plan	RGB, HSV	Linear Logistic Regression
CHEN et collab. [2014]	SIFT (dense)	Moment couleur	RGB	SVM
SIXTA [2011]	MBLBP	—	—	k-NN
ŠULC et MATAS [2013, 2017]	MSLBP	—	—	SVM
SULC et MATAS [2014]	Ffirst	—	—	SVM

L'analyse de la texture est un problème bien posé dans le domaine de la reconnaissance d'objets. Elle regroupe différents domaines d'étude, à savoir, la classification, la segmentation, la synthèse, et la compression de la texture TUCERYAN et JAIN [1993]. Un intérêt particulier était porté à la classification de la texture durant les deux dernières décennies LIU et collab. [2019]. La classification de la texture est généralement composée de deux phases : phase d'extraction et représentation des attributs, et phase de classification. Néanmoins, la plupart des travaux sur la classification de la texture se sont concentrés sur la première phase de description et caractérisation étant la plus importante dans le processus de classification de la texture. En effet, si la représentation de la texture n'est pas bien établie, les classifieurs les plus performants auront des performances relativement faibles.

Par conséquent, les précédents travaux réalisés dans le cadre de la classification des espèces d'écorce ont appliqué des méthodes de représentation de texture pour extraire et encoder les attributs de texture caractéristiques de la surface d'écorce. Dans ce qui suit, nous présentons un aperçu de ces travaux qu'on peut regrouper selon la mise à échelle ou non de la représentation, i.e., le descripteur final est un descripteur de structures locales ou de structures multi-échelles.

Par ailleurs, la couleur de l'écorce est le second caractère visuel utilisé par les botanistes afin d'identifier certaines espèces. En effet, on peut distinguer des écorces ayant visuellement la même

0	0	0	1
1	1	1	1
2	2	2	3
3	3	4	5

(a)

	0	1	2	3	4	5
0	2	1	0	0	0	0
1	1	3	0	0	0	0
2	0	0	2	1	0	0
3	0	0	1	1	1	0
4	0	0	0	1	0	1
5	0	0	0	0	1	0

(b)

FIGURE 2.1 – Exemple de calcul de la matrice de cooccurrences. (a) Image en entrée. (b) Matrice de cooccurrences pour un déplacement $(d, \theta) = (1, 0^\circ)$ TUCERYAN et JAIN [1993].

apparence (texture) mais avec une différence de couleurs de la surface (grisâtre, rougeâtre, brunâtre, jaunâtre, etc), la couleur des lenticelles (blanc, noir), la couleur entre les fissures et les découpes, et aussi la différence d'intensité de la couleur entre les branches et la base de l'arbre (un peu plus claire ou plus foncée) WOJTECH [2011]. Pour ce qui est de l'ajout de l'information contenue dans la couleur, nous présentons également l'ensemble des travaux ayant ajouté cette primitive pour représenter la surface de l'écorce, le plus souvent en complémentarité avec la représentation de la texture. Le Tableau 2.1 résume les travaux réalisés pour la classification de l'écorce.

2.1.1 Représentation de la texture d'écorce

2.1.1.1 Représentation spatiale locale

2.1.1.1.1 Matrice de cooccurrences

Les premiers travaux sur la description de la texture d'écorce étaient basés sur la modélisation de la texture par des statistiques du k^{eme} ordre, i.e., les statistiques d'occurrence des intensités de pixels. L'approche la plus répandue dans cette perspective est celle des matrices de cooccurrences (*Gray level co-occurrence matrices (GLCM)* HARALICK et collab. [1973]).

Une matrice de cooccurrences décrit les statistiques de second ordre de l'image. Soit l'image en entrée \mathcal{I} et une paire de pixels aux coordonnées (h, w) et (h', w') ayant respectivement les niveaux d'intensité x_i, x_j . Ces pixels sont spatialement répartis selon une distance d et une orientation θ .

La matrice de cooccurrences $M_{d,\theta}$ est de taille $G \times G$ où G est le niveau de quantification de l'intensité (généralement, $G = 256$ niveaux de gris). $M_{d,\theta}$ est définie comme suit :

$$M_{d\theta}(x_i, x_j) = |((h, w), (h', w'))|; \quad I(h, w) = x_i, I(h', w') = x_j \quad (2.1)$$

où $|\cdot|$ est l'opérateur de cardinalité. $M_{d,\theta}(x_i, x_j)$ est la fréquence d'apparition de la paire $((h, w), (h', w'))$ selon le déplacement d et l'orientation θ dans l'image \mathcal{I} . La Figure 2.1 illustre un exemple de calcul de la matrice de cooccurrences pour une image quantifiée à $G = 5$ niveaux de gris, où $d = 1$ et $\theta = 0^\circ$.

De nombreux attributs statistiques peuvent être dérivés de la matrice de cooccurrences pour

représenter les caractéristiques de la texture et former ainsi le vecteur descripteur de l'image. Ces attributs, appelés statistiques de second ordre, sont proposés par Haralick [HARALICK et collab. \[1973\]](#). Les plus fréquemment utilisés sont l'énergie, l'entropie, le contraste, l'homogénéité, et la corrélation. Les attributs sont définis par les formules suivantes :

$$\begin{aligned}
 \text{Energie} &= \sum_{x_i, x_j}^G M_{d,\theta}(x_i, x_j)^2 \\
 \text{Entropie} &= - \sum_{x_i, x_j}^G M_{d,\theta}(x_i, x_j) \log(M_{d,\theta}(x_i, x_j)) \\
 \text{Contraste} &= \sum_{x_i, x_j}^G (x_i - x_j)^2 M_{d,\theta}(x_i, x_j) \\
 \text{Homogénéité} &= \sum_{x_i, x_j}^G \frac{M_{d,\theta}(x_i, x_j)}{1 + |x_i - x_j|} \\
 \text{Corrélation} &= \sum_{x_i, x_j}^G \frac{(x_i - \mu_h)(x_j - \mu_w) M_{d,\theta}(x_i, x_j)}{\sigma_h \sigma_w}
 \end{aligned} \tag{2.2}$$

μ_h, μ_w sont les moyennes et σ_h, σ_w sont les écarts-types de $M_{d,\theta}(G_h)$ et $M_{d,\theta}(G_w)$ respectivement (selon les lignes G_h et les colonnes G_w de la matrice de cooccurrences) tel que :

$$M_{d,\theta}(G_h) = \sum_{x_i} M_{d,\theta}(G_h, x_i) \text{ et } M_{d,\theta}(G_w) = \sum_{x_j} M_{d,\theta}(G_w, x_j).$$

On retrouve parmi les travaux de classification d'espèces d'écorce ceux basés sur des approches statistiques comme [WAN et collab. \[2004\]](#) qui a établi une comparaison entre différentes approches statistiques pour la description des images d'écorce, à savoir : la matrice de cooccurrences (GLCM), la méthode d'histogramme, la méthode d'auto-corrélation, et le codage par plage (*Run-Length Method*). Les meilleures performances ont été rapportées par les GLCM pour $d = 1$ et $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$, et avec un vecteur descripteur de la texture composé de dix attributs de Haralick.

Les GLCM ont aussi été utilisées dans [BRESSANE et collab. \[2016\]](#); [LAMEI et collab. \[2010\]](#) pour la description de l'image de texture de l'écorce avec les cinq attributs de Haralick cités dans 2.2. Les auteurs [BRESSANE et collab. \[2015b\]](#) proposent un descripteur de la texture qui combine les statistiques extraites de la GLCM et celles extraites de l'histogramme de l'image [BRESSANE et collab. \[2015a\]](#), et qui sont : l'uniformité (l'énergie), l'entropie, l'asymétrie (skewness), et la régularité (smoothness).

D'autre part, [HUANG et collab. \[2006d\]](#) propose de combiner les attributs des GLCM calculés avec $d = 4$ et $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ avec la dimension fractale pour représenter la complexité de la texture d'écorce à différentes échelles.

Les matrices de cooccurrences ont prouvé leur efficacité pour représenter la structure locale de la texture d'écorce avec un simple calcul. Cependant, plus le niveau de quantification d'intensité de l'image en entrée est élevé, plus le temps de calcul et le stockage en mémoire est considérable. D'autant plus qu'on génère autant de GLCM pour chacune des configurations (d, θ) désirées pour différentes représentations et orientations de la structure de la texture.

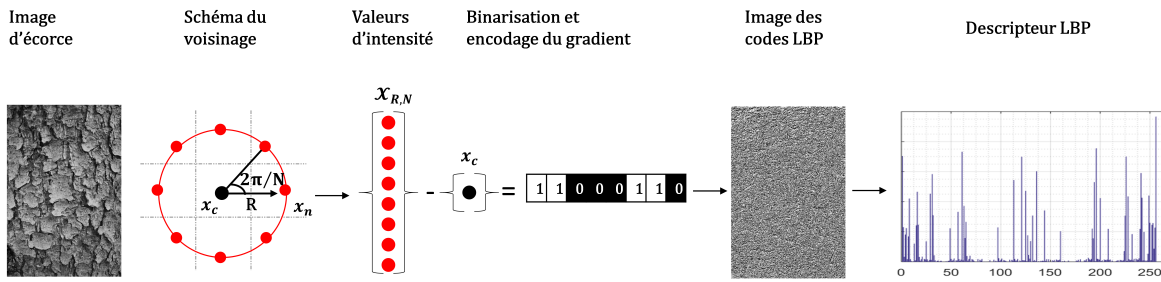


FIGURE 2.2 – Les étapes de génération du descripteur LBP.

2.1.1.1.2 Descripteur LBP

Autre que les matrices de cooccurrences pour décrire la structure de la texture d'écorce, on trouve le descripteur des micro-motifs binaires locaux (*Local Binary Pattern*(LBP)). Introduit initialement par [OJALA et collab. \[1996\]](#), le LBP encode les microstructures de la texture comme les coins, les contours, et les bords.

Le principe du descripteur LBP repose sur la représentation d'un pixel donné dans l'image par son voisinage en encodant les différences d'intensité. Spécifiquement, le LBP encode le signe de la variation d'intensité entre un pixel donné x_c dans l'image, appelé pixel central, et ses N voisins équitablement répartis sur un cercle de rayon R . Un code binaire est construit de sorte que si la valeur d'intensité d'un voisin x_n est supérieure ou égale à celle du pixel central x_c on attribut 1, sinon 0. Ensuite, le code décimale calculé à partir de cette chaîne binaire représente le code LBP du motif local. Formellement, un code LBP est défini comme suit :

$$\text{LBP}_{R,N} = \sum_{n=0}^{N-1} f(x_n - x_c)2^n; \quad f(d) = \begin{cases} 1 & \text{si } d \geq 0 \\ \text{sinon } 0. \end{cases} \quad (2.3)$$

Cet encodage est réalisé pour tous les pixels de l'image en entrée, produisant une image de code LBP. La distribution de ces codes est représentée par un histogramme de 2^N bins formant le descripteur LBP. La Figure 2.2 montre un exemple de calcul du descripteur LBP avec la configuration $(R, N) = (1, 8)$.

Le descripteur LBP a la propriété d'être simple à calculer et invariant aux changements monotoniques du niveau de gris. Cependant, le principe de calcul du descripteur LBP le rend sensible au bruit, i.e., la moindre fluctuation de la valeur d'intensité du pixel central x_c (le seuil de binarisation) engendre un code LBP différent. De plus, il a la limite de ne pas être invariable aux changements de rotation locale, i.e., le principe de calcul du gradient orienté est fortement lié à l'emplacement spatial des voisins autour de x_c , une rotation du voisinage d'un angle donné, affecte principalement les valeurs des poids qui produisent à leur tour un code LBP différent.

Par conséquent, [OJALA et collab. \[2002\]](#) a introduit des extensions du descripteur LBP, à savoir, le principe d'uniformité afin de remédier aux problèmes de sensibilité au bruit, l'invariance à la rotation, et la représentation multi-résolutions.

Selon la transition spatiale dans un code binaire LBP, [OJALA et collab. \[2002\]](#) distingue deux types de motifs : motifs LBP uniformes, et motifs LBP non-uniformes. Un code binaire LBP est dit uniforme si le nombre de transitions entre 1/0 et vis vers ça dans la chaîne binaire ne dépasse pas

une mesure d'uniformité $U = 2$, autrement, il est dit non-uniforme. La mesure d'uniformité U est donnée par :

$$U(\text{LBP}_{R,N}) = \sum_{n=0}^{N-1} |f(x_n - x_c) - f(x_{\text{mod}(n+1,N)} - x_c)| \quad (2.4)$$

Les motifs uniformes représentent peu de variations dans les valeurs d'intensité dans le voisinage, tandis que les motifs non-uniformes représentent une grande variation, généralement due à un bruit.

Appliquer cette configuration, ou bien *mapping*, dénommée $u2$, revient à attribuer les motifs uniformes à des bins distincts et regrouper tous les motifs non-uniformes dans le même bin dans l'histogramme LBP. Ce mapping résulte en un descripteur $\text{LBP}_{R,N}^{u2}$ de dimension réduite égale à $N(N-1) + 3$. De plus, OJALA et collab. [2002] a observé que les motifs uniformes caractérisent des propriétés fondamentales de la texture.

Afin d'assurer une invariance à la rotation, OJALA et collab. [2002] propose le mapping *Rotation invariant (ri)* qui effectue un décalage circulaire bit par bit vers la droite dans la chaîne binaire $\text{LBP}_{R,N}$ n fois. Formellement, $\text{LBP}_{R,N}^{ri}$ est donné par :

$$\text{LBP}_{R,N}^{ri} = \min\{\text{ROR}(\text{LBP}_{R,N}, n) \mid n = 0, 1, \dots, N\} \quad (2.5)$$

De plus, OJALA et collab. [2002] combine les deux précédents mappings pour une invariabilité au bruit et aux changements de rotation locale, en définissant le mapping *Rotation Invariant Uniform 2, riu2* donné par :

$$\text{LBP}_{R,N}^{riu2} = \begin{cases} \sum_{n=0}^{N-1} f(x_n - x_c) & \text{si } U \leq 2 \\ \text{sinon} & N + 1 \end{cases} \quad (2.6)$$

Ce mapping permet de réduire d'avantage la dimension du descripteur $\text{LBP}_{R,N}^{riu2}$ à $N + 2$ bins.

Parmi les travaux réalisés dans le contexte de l'identification et la classification des espèces d'écorce avec le LBP, on retrouve BLANCO et collab. [2016] qui propose une approche mobile pour l'identification des espèces d'écorce en caractérisant sa texture par un LBP basique. Une étape de prétraitement a précédé la description de la texture, elle consiste en l'extraction du fond et la normalisation de la luminance par un filtre moyen.

La version invariable à la rotation LBP^{ri} est testée pour la description des micro-motifs de l'écorce à ImageCLEF 2013 BAKIĆ et collab. [2013] et PlantCLEF 2014 GOËAU et collab. [2014b]. Par ailleurs, lors de la campagne d'évolution PlantCLEF 2015, MICA LE et collab. [2015] propose l'utilisation des descripteurs à base de noyaux (Kernel Descriptor) BO et collab. [2010] utilisant une représentation LBP.

Le LBP encode les micro-motifs (ou bien les microstructures) de la texture, comme les coins, les contours, et les bords. Cependant, la texture de l'écorce est formée de grandes structures, dites macrostructures. Dans la suite, nous présentons les différents travaux sur la représentation de la texture d'écorce à base de descripteurs multi-échelles, qui permettent d'encoder l'information macrostructure.

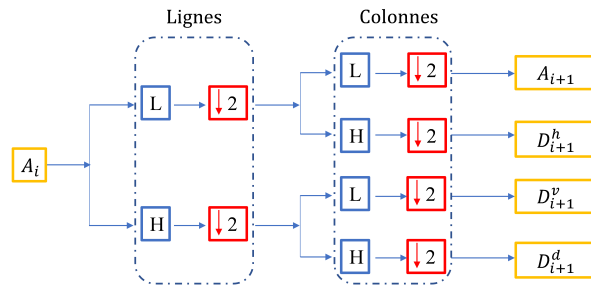


FIGURE 2.3 – Décomposition en ondelettes (2DDWT).

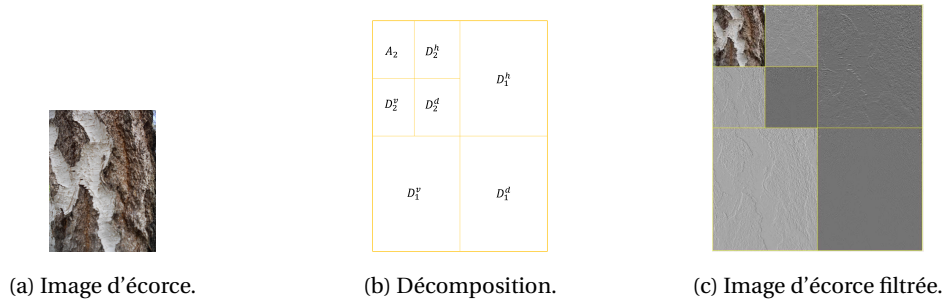


FIGURE 2.4 – Décomposition en ondelettes d'une image d'écorce en 2 niveaux.

2.1.1.2 Représentation spatiale multi-échelles

Pour une représentation multi-échelles de la texture, il convient souvent d'appliquer des techniques de traitement de signal, i.e. des filtres spatio-fréquentiels, comme les ondelettes et leurs variantes telles que les filtres de Gabor. Ces approches permettent une analyse multi-résolutions et multi-orientations de la structure. En plus de ces approches, on trouve des travaux basés sur la description des points d'intérêt couplée à la technique du sac à mots (*Bag of Words(BoW)*) pour générer le descripteur final de la texture, ou plus récemment, l'application de différentes variantes LBP multi-échelles. Ces approches seront revues dans ce qui suit.

2.1.1.2.1 Filtres spatio-fréquentiels

La décomposition en ondelettes discrètes 2D (*Discrete Wavelet Transform(DWT)*) [MALLAT \[1989\]](#) est réalisée par l'application de deux types de bancs de filtres : un filtre passe-bas Lp et un filtre passe-haut Hp sur l'image en entrée. Cette opération, dite convolution, est d'abord réalisée sur les lignes, ensuite sur les colonnes. Après un sous-échantillonnage, l'image sera représentée par quatre coefficients : une *approximation* (A_1) et trois *détails* D_1^h , D_1^v , D_1^d qui permettent la détection de contours selon trois directions : horizontale, verticale, et diagonale. La Figure 2.3 montre l'application des filtres Lp , Hp et la générations des coefficients A_1 , D_1^h , D_1^v , D_1^d .

Pour une analyse multi-résolutions, une pyramide de décomposition en ondelettes est construite, où, à chaque niveau i , on reprend la convolution de l'image d'approximation A_{i-1} avec les bancs de filtres Lp et Hp . La Figure 2.4 illustre une décomposition d'une image d'écorce en deux niveaux.

La distribution de l'information contenue dans chaque image dans la pyramide de décomposition est approximée par des attributs statistiques. Les attributs les plus utilisés sont la moyenne et l'écart type, on retrouve également dans certains travaux l'énergie [HUANG \[2006\]](#); [KIM et collab. \[2011\]](#).

Dans [SONG et collab. \[2004\]](#), les auteurs introduisent une représentation des structures d'écorce et leur distribution. Les attributs relatifs à la longueur et l'orientation de la structure sont extraits à partir d'une représentation binaire de l'image d'écorce. Sur l'image en entrée, [SONG et collab. \[2004\]](#) applique une série d'ondelettes, ensuite, l'image moyenne des approximations est segmentée afin de produire une image binaire des structures d'écorce. Pour représenter la longueur de la structure, les auteurs gardent pour chaque pixels le nombre maximum de pixels connectés à celui là dans les quatre directions principales, donnant le descripteur *Long Connection Length Emphasis (LCLE)*. De plus, [SONG et collab. \[2004\]](#) combine le LCLE avec l'attribut de l'énergie extrait de la GLCM calculée à partir de l'image en niveau de gris. Le descripteur final est une concaténation de ces deux représentations qui encodent respectivement, l'orientation et la longueur des structures à travers différentes échelles, et leur distribution statistique de second ordre.

Cependant, malgré l'efficacité des DWT pour une représentation multi-résolutions, leur utilisation est limitée par la variation à la translation et la détection de contours en seulement trois directions, qui est parfois insuffisant pour représenter les contours des images naturelles. Par conséquent, d'autres variantes d'ondelettes ont été étudiées pour la représentation de la surface de l'écorce.

Les filtres de Gabor (une fonction d'ondelette) sont parmi les descripteurs de texture les plus connus, on reconnaît l'efficacité des filtres de Gabor pour la représentation des images de texture de manière similaire au système visuel humain [MANJUNATH et MA \[1996\]](#). Formellement, un filtre de Gabor est une représentation complexe définie par le produit d'une sinusoïde orientée et une enveloppe Gaussienne. La partie réelle *Gabor*(h, w) est définie par la formule suivante :

$$Gabor(h, w) = \exp\left(-\frac{h'^2 + w'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{h'}{\lambda}\right) \quad (2.7)$$

où :

$$h' = h \cos(\theta) + w \sin \theta$$

$$w' = -h \sin(\theta) + w \cos \theta$$

où : λ représente la longueur de l'onde sinusoïdale, i.e., l'échelle, θ représente l'orientation du filtre, et σ représente l'écart type de l'enveloppe Gaussienne.

Plus la valeur de σ est élevée, plus l'image est floutée et les détails locaux fins sont estompés.

Les filtres de Gabor on été utilisés pour la description de la texture de l'écorce dans les travaux de [CHI et collab. \[2003\]](#); [HUANG et collab. \[2006a,b\]](#). Le descripteur de texture combine des attributs statistiques extraits des images réponses, à savoir, la moyenne et l'écart type.

Les auteurs dans [BERTRAND et collab. \[2017\]](#) ont utilisé les filtres de Gabor pour encoder l'orientation des structures de l'écorce par des statistiques extraites à partir des images réponses, comme réalisé dans [HUANG et collab. \[2006a\]](#). Cependant, [BERTRAND et collab. \[2017\]](#) utilisent seulement deux orientations, 0° et 90° , et 4 échelles, où une image moyenne des quarts échelles est retenue pour une direction donnée. Le vecteur descripteur est formé par la concaténation des mesures statistiques globales, moyenne et écart type, extraites des deux images moyennes à l'orientation 0° et 90° . En plus, ce descripteur est combiné à une représentation de l'orientation et de la forme des structures assez similaire à [SONG et collab. \[2004\]](#). Tout d'abord, une image binaire est générée suite à l'application du détecteur de contours Canny [CANNY \[1987\]](#). Ensuite la distribution

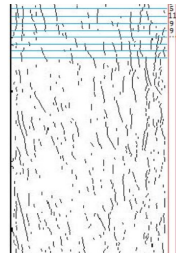


FIGURE 2.5 – Exemple de calcul du vecteur d'intersections BERTRAND et collab. [2017].

des structures est donnée par la somme des nombres d'intersections des points avec les lignes d'une grille superposée sur l'image de contours verticaux, donnant le vecteur des mots verticaux. Ce même encodage de la distribution est réalisé sur l'image des contours horizontaux suivant les colonnes de la grille, pour le calcul du vecteur des mots horizontaux. La représentation finale est la concaténation des deux mots, nommée Gw . la Figure 2.5 montre un exemple du Gw .

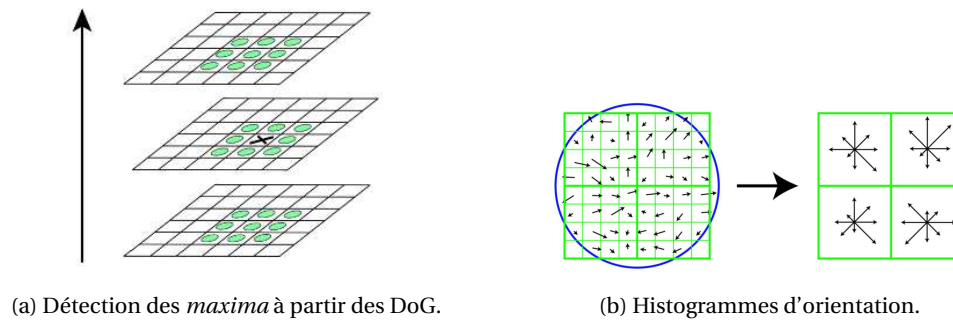
Une autre variante d'ondelettes similaire aux filtres de Gabor avec une invariance à la translation, est la représentation en ondelettes complexes. Introduites par KINGSBURY [2001], les ondelettes complexes sont représentées par un arbre dual avec partie réelle et partie imaginaire (*Dual Tree Complex Wavelet Transform*(DT-CWT)). TSCHIEDEL GUEDES et BORGES [2018] a appliqué les DT-CWT pour la description de diverses structures d'écorce. Les caractéristiques retenues étaient la moyenne et l'écart type extraites à partir des coefficients de la DT-CWT pour quatre échelles et six orientations.

HUANG et collab. [2006c] a testé les performances des contourlets (*contourlet transform*(CT)) pour la représentation de la texture de l'écorce. En effet, les CT DO et VETTERLI [2005] sont un type d'ondelettes qui permet de représenter les contours fins et lisses présents dans les images naturelles grâce à l'utilisation d'une représentation multi-échelles avec une pyramide Laplacien, et l'application d'un ensemble de filtres directionnels. OTHMANI et collab. [2013] a proposé une représentation des images 2D de texture d'écorce basée sur la combinaison des ondelettes, notamment les contourlets (CT), les ondelettes complexes (DT-CWT), et une variante directionnelle appelée Rotated Complex Wavelet Filters (DT-RCWF). Les images 2D sont reconstruites à partir d'un nuage de points 3D de la surface de différents types d'écorce.

L'utilisation de différents types d'approches à base de filtres spatio-fréquentiels était en particulier motivée par les performances reportées en comparaison avec les GLCM, d'autant plus par la description compacte de l'information provenant d'une analyse multi-échelles et multi-orientations grâce à l'utilisation des attributs statistiques. Cependant, la génération des bancs de filtres n'est pas en elle même coûteuse en temps de calcul, mais dépend du domaine d'application. De plus, l'opération de convolution avec ces bancs de filtre (ou bien le produit scalaire dans l'espace fréquentiel) qui devient d'une complexité et d'un temps de calcul notable pour des images de haute résolution.

2.1.1.2.2 Points d'intérêt et BoW

Une représentation par points d'intérêt fait référence à une description locale des images, car encode l'information présente dans certaines régions de manière adaptative au contenu. Le plus

FIGURE 2.6 – Détecteur et descripteur de points d'intérêt SIFT [LOWE \[2004\]](#).

souvent, la détermination de ces régions passe par une phase de détection des points d'intérêt. Ces derniers représentent les structures les plus fréquentes où il y a des changements conséquents d'intensité. Ensuite, pour chaque point d'intérêt détecté, une description locale est calculée à partir de l'information contenue dans la région autour de lui.

L'objectif de la description par points d'intérêt est de permettre l'encodage de l'information pertinente à partir de régions spécifiques sans passer par une phase de segmentation pour isoler ces régions.

Initialement, la description par points d'intérêt était utilisée pour la mise en correspondance des objets similaires pris sous différents angles de vue. Plus tard, cette approche a été intégrée dans des applications d'analyse de texture [ZHANG et collab. \[2007\]](#). Différents détecteurs et descripteurs de points ont été proposés, comme SIFT [LOWE \[2004\]](#), SURF [BAY et collab. \[2006\]](#), détecteur Harris pour les régions affines [MIKOLAJCZYK et SCHMID \[2002\]](#). Une revue détaillée de ces détecteurs/descripteurs est présentée dans [MIKOLAJCZYK et SCHMID \[2005\]](#)

Parmi les approches de détection et description des points d'intérêt les plus utilisées, on trouve la transformée invariante à l'échelle (*Scale-Invariant Feature Transform* (SIFT)) [LOWE \[2004\]](#) (cf Figure 2.6). Les points SIFT sont d'abord détectés comme des points *maxima* dans un espace d'échelle formé par une pyramide de différences Gaussiennes (*Difference of Gaussian* DoG) afin d'assurer une propriété d'invariance aux changements d'échelle. Un exemple de détection d'un point SIFT à partir d'une DoG est donné dans la Figure 2.6a. Ensuite, un histogramme d'orientations est calculé à partir d'une fenêtre Gaussienne centrée autour de ce point d'intérêt, et dont le rayon est proportionnel à l'échelle du point d'intérêt (le niveau de DoG ou ce point a été détecté). La combinaison des histogrammes d'orientation normalisés comme illustré dans la Figure 2.6b forme un descripteur SIFT d'une taille fixe égale à 128. La description finale d'une image de texture est donc une matrice 2D dont la hauteur est égale au nombre de points d'intérêt détectés dans cette image.

Ainsi, une description par points d'intérêt génère un descripteur par point détecté. Toutefois, il convient de produire un seul descripteur global pour caractériser l'image de texture. Pour cela, le modèle de sac à mots (*Bag of Words* BoW) est utilisé. Le BoW était à la base introduit pour le traitement du langage naturel, et a été intégré pour la description du contenu visuel des images, donnant ainsi le modèle de sac à mots visuels (*Bag of Keypoints*) [CSURKA et collab. \[2004\]](#). Le principe ici consiste à apprendre un dictionnaire à partir des descriptions locales des régions autour des points d'intérêt, ou bien un pixel, afin de générer une représentation globale des occurrences

des mots visuels pour un ensemble d'images données.

Pour se faire, on procède par un encodage par vote. La construction du dictionnaire des mots visuels passe par une étape de clustering (avec *k-means*) des descripteurs locaux en un nombre prédéterminé de cluster (qui donne aussi la taille du dictionnaire), et la génération des mots visuels qui sont les centres des clusters. Chaque descripteur local dans l'image est ensuite associé à un cluster (c'est à dire à un mot du dictionnaire). Au final l'histogramme des fréquences de ces mots de taille égale à celle du dictionnaire représente ainsi le descripteur globale de l'image.

Dans cette catégorie, on trouve le travail de [WENDEL et collab. \[2011\]](#) pour l'identification des espèce végétales à partir des images de feuilles, d'aiguilles, et aussi d'écorce. [WENDEL et collab. \[2011\]](#) a appliqué une description locale par points d'intérêt SIFT sur des images d'écorce. Une description globale était dérivée suite à l'utilisation du modèle BoW sur l'ensemble des descripteurs SIFT pour générer un descripteur global.

On retrouve aussi une combinaison entre différents détecteurs et descripteurs de points d'intérêt, comme réalisé par PlantNet, [BAKIĆ et collab. \[2013\]](#) avec le détecteur des coins Harris [HARRIS et STEPHENS \[1988\]](#) et le descripteur SURF [BAY et collab. \[2006\]](#). Ce descripteur suit le même principe de la pondération par une fenêtre Gaussienne que SIFT, mais basé sur les réponses des ondelettes de Haar dans cette fenêtre. Une représentation similaire était testée dans l'édition de 2014 de PlantCLEF avec une analyse multi-résolutions et multi-orientations lors de la phase de détection des points d'intérêt avec Harris [GOËAU et collab. \[2014b\]](#).

Il existe une extension de la représentation locale à base des points d'intérêt vers une représentation dite 'dense' afin d'extraire plus d'attributs et d'encoder plus de détails de la texture. Le principe de cette représentation dense est de considérer chaque pixel dans l'image comme un point d'intérêt, et de procéder à la description de la région autour de lui. Ainsi, une représentation globale peut être générée en utilisant le principe de BoW. Cette représentation était évaluée par NLab UTokyo [NAKAYAMA \[2013\]](#) lors de ImageCLEF 2013, et lors de la campagne PlantCLEF 2014 par IBM AU [CHEN et collab. \[2014\]](#) en utilisant les points d'intérêt SIFT.

La représentation par points d'intérêt offre des propriétés d'invariance aux changements d'échelle, et de rotation, suite au processus de la détection et la description. Cependant, c'est une représentation parcimonieuse de la surface de la texture. Par conséquent, elle peut ne pas coïncider avec des éléments importants de la texture, et provoque ainsi une perte d'information. De plus, pour des images de taille réduite, les régions détectées peuvent ne pas être suffisantes pour représenter le contenu visuel. De ce fait, on trouve de nombreux travaux qui se sont penchés sur une représentation dense du contenu visuel, i.e., en considération tous les pixels de l'image.

La phase d'apprentissage pour la génération du dictionnaire des mots visuels, dans BoW, permet une interprétation plus générale et cohérente du contenu visuel à partir d'une description locale. Cependant, ceci requiert une base avec des classes d'images variées afin de pouvoir généraliser cette description sur des nouveaux échantillons, faute de ça, le réapprentissage d'un nouveau dictionnaire est appliqué pour chaque nouvelle classe dans la base d'images afin de générer son propre modèle de textons. De plus, la génération du dictionnaire des textons à travers une phase de clustering par *k-means* devient de plus en plus coûteuse pour des dictionnaires de textons de taille élevée. En effet, plus la taille du dictionnaire est élevée (donc plus de textons), plus

les structures de la texture sont mieux caractérisées.

2.1.1.2.3 LBP multi-échelles

Nous passons maintenant aux approches à base de micro-motifs binaires LBP en représentation multi-échelles. En effet, le LBP basique avec son support spatial très local de 3×3 est limité à encoder seulement les micro-motifs, et ne permet pas de représenter de larges structures.

Afin de permettre l'encodage et la représentation des macrostructures, [OJALA et collab. \[2002\]](#) a introduit une version LBP multi-résolutions, qu'on dénomme MResLBP, en combinant des descripteurs $LBP_{R,N}$ avec des configurations différentes de (R, N) . Spécifiquement, MResLBP est formé de la concaténation (\oplus) de descripteurs $LBP_{R,N}$ avec des valeurs graduellement élevées de rayons R et de voisins N . MResLBP est formalisé comme suit :

$$MResLBP_{(1,8)+(2,16)+(3,24)}^{riu2} = LBP_{1,8}^{riu2} \oplus LBP_{2,16}^{riu2} \oplus LBP_{3,24}^{riu2} \quad (2.8)$$

La représentation de MResLBP permet d'encoder des structures grossières étendues sur un large support spatial. Par conséquent, ce schéma a amélioré les performances de LBP, et a été fortement utilisé dans de multiples variantes LBP pour assurer une représentation multi-échelles. Des synthèses sur ces variantes sont présentées dans [HUANG et collab. \[2011\]](#); [KE-CHEN et collab. \[2013\]](#); [LIU et collab. \[2017\]](#); [NANNI et collab. \[2012\]](#); [ZHAO \[2011\]](#).

Parmi les variantes LBP multi-échelles utilisées pour la texture d'écorce, on cite Multi-Block LBP (MBLBP) [SIXTA \[2011\]](#) basée sur une représentation dense de l'espace d'échelle de [LIAO et collab. \[2007\]](#), Multi-Scale LBP (MSLBP) [ŠULC et MATAS \[2013\]](#); [SULC et MATAS \[2014\]](#); [ŠULC et MATAS \[2017\]](#) inspiré de LBP Filtering (LBPf) [MÄENPÄÄ et PIETIKÄINEN \[2003\]](#), et de Pyramid-based LBP (PLBP) [QIAN et collab. \[2011\]](#).

Ces approches se distinguent entre elles par rapport à : la définition et la construction de l'espace d'échelle, le filtre passe-bas utilisé, et enfin l'encodage de l'information d'intensité dans la macrostructure. Dans ce qui suit une description détaillée est présentée pour chacune de ces variantes.

Multi-Block LBP (MBLBP) : pour le calcul du descripteur MBLBP [LIAO et collab. \[2007\]](#), un filtre moyen avec des tailles de fenêtres variables en fonction du niveau de l'espace d'échelle est appliqué à l'image en entrée. La taille du filtre MBLBP doit être un multiple de 3 pour pouvoir construire un voisinage de 9 blocs, dont le bloc central. Multi-Block LBP (MBLBP) est ainsi basé sur l'encodage des valeurs des moyennes des blocs de pixels, au lieu des simples valeurs d'intensité des pixels du voisinage. Le calcul des moyennes de blocs peut être accéléré par l'utilisation des images intégrales [VIOLA et JONES \[2001\]](#). La Figure 2.7 montre un exemple du filtre MBLBP de 9×9 . Après le calcul des moyennes locales de chaque bloc, un encodage basique de LBP est appliqué à chaque niveaux pour représenter le motif local.

[LIAO et collab. \[2007\]](#) a également proposé une mesure d'uniformité pour réduire la taille du descripteur final. Une valeur de seuil est prédéfinie pour déterminer les motifs uniformes selon leurs fréquences dans l'histogramme MBLBP. Les 63 motifs les plus fréquents (ayant des valeurs de bins élevées après avoir trier dans l'ordre décroissant l'histogramme MBLBP résultant) sont appelés uniformes, le reste des motifs est attribué au même bin et sont appelés motifs non-uniformes.

SIXTA [2011] a repris la représentation multi-block de **LIAO et collab. [2007]** pour la représentation des écorces. Cependant, **SIXTA [2011]** introduit une représentation dense de l'espace d'échelle avec 10 échelles à partir de 9×9 avec un facteur d'incrément de 1.291. Le descripteur final, est ainsi une concaténation de dix histogrammes LBP.

LBP Filtering (LBPF) : MÄENPÄÄ et PIETIKÄINEN [2003] combine un voisinage multi-échelles avec une série de filtres Gaussiens. Les niveaux de l'espace d'échelle suivent une distribution exponentielle, et le filtre Gaussien permet de calculer une moyenne gaussienne pour une région autour d'un point voisin donné. Ces régions sont appelées des "régions effectives" **MÄENPÄÄ et PIETIKÄINEN [2003]**, et sont schématisées dans La Figure 2.8a par des cercles répartis équitablement dans le voisinage. Les rayons r_i des régions effectives sont donnés par la formule suivante :

$$r_i = r_{i-1} \left(\frac{2}{1 - \sin(\pi/N_i)} - 1 \right), \quad i \in \llbracket 2, \mathcal{L} \rrbracket, \quad r_1 = 1.5 \quad (2.9)$$

Par conséquent, les valeurs des niveaux d'échelle (ou bien des rayons LBP à chaque niveau d'échelle), R_i sont calculées de sorte que les cercles des régions effectives se touchent. La Figure 2.8a illustre ces rayons par des cercles pointillés. Les valeurs de R_i sont calculées comme suit :

$$R_i = \frac{r_i + r_{i-1}}{2} \quad (2.10)$$

Un descripteur $LBP(R_i, 8)$ est calculé pour chaque niveau de l'espace d'échelle. Le descripteur final LBPF est formé par la concaténation des descripteurs LBP comme illustré par la Figure 2.8b.

Multi-Scale LBP (MSLBP) : pour le descripteur MSLBP **ŠULC et MATAS [2013, 2017]**, le schéma de l'espace d'échelle est très similaire à celui de LBPF concernant l'utilisation du filtre Gaussien pour déterminer les régions effectives. Cependant, la différence réside dans la distribution des niveaux d'échelle (ou bien les rayons R_i à chaque niveau). Ces derniers sont répartis selon une distance de déplacement radiale égale à $\sqrt{2}$ faisant que les cercles des régions effectives se recouvrent dans le sens radial, comme le montre la Figure 2.9a. Les valeurs des R_i sont données par la formule suivante :

$$R_i = R_{i-1} \sqrt{2} \quad (2.11)$$

Cette représentation dense de l'espace d'échelle permet de représenter des détails fins et encoder d'avantage l'information structurale. De plus, lors de la génération du descripteur final, les descripteurs LBP calculés par échelle sont concaténés en effectuant un décalage d'un niveau d'échelle donnant un descripteur multi-échelles multiple, comme illustré dans la Figure 2.9b. Cette redondance de représentation vise à pondérer la contribution des histogrammes LBP très locaux dans la représentation finale de l'image de texture.

Pyramid-based LBP (PLBP) : le descripteur PLBP **QIAN et collab. [2011]** se présente comme une généralisation des descripteurs LBPF et MSLBP. Le principe de PLBP consiste à effectuer une représentation multi-échelles de la texture par une décomposition en pyramide Gaussienne (cf. Figure 2.10a) afin d'encoder les structures de texture pouvant figurer à différentes résolutions. À chaque niveau de la pyramide Gaussienne, un filtre Gaussien est appliqué pour déterminer

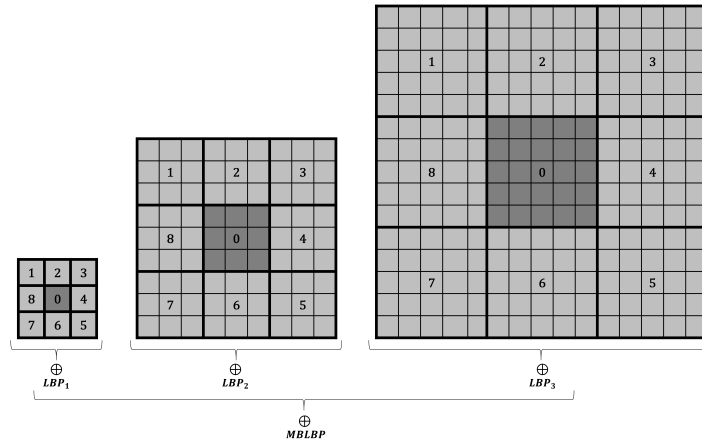


FIGURE 2.7 – Espace d'échelle de MBLBP LIAO et collab. [2007], et construction du descripteur multi-échelles.

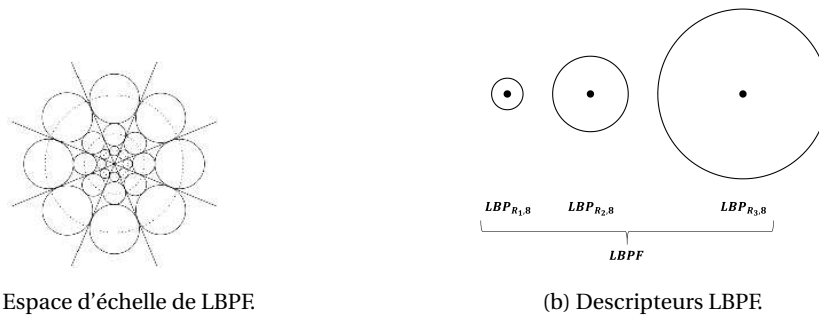


FIGURE 2.8 – La variante multi-échelles LBPF (image (a) prise de MÄENPÄÄ et PIETIKÄINEN [2003]).

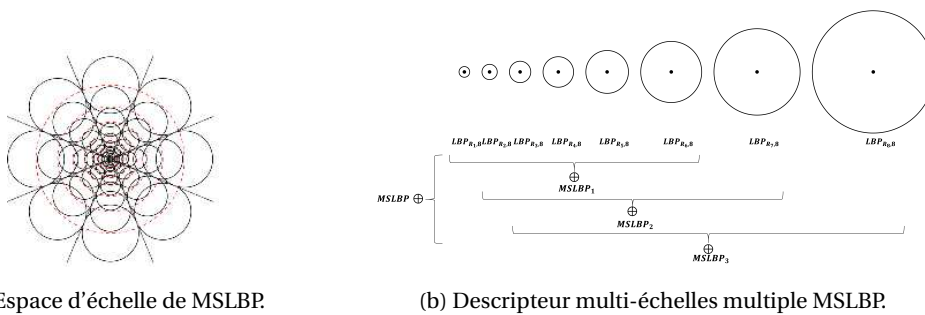


FIGURE 2.9 – La variante multi-échelles MSLBP (image (a) prise de MÄENPÄÄ et PIETIKÄINEN [2003]).

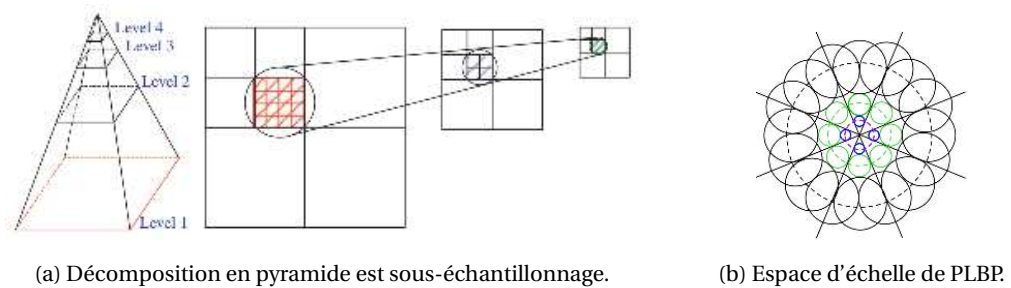


FIGURE 2.10 – La variante multi-échelles PLBP (images prises de QIAN et collab. [2011]).

les régions effectives, avec éventuellement, un nombre variable de points voisins (d'où le recouvrement des cercles dans le sens circonférentiel dans la Figure 2.10b), ensuite, l'image est sous-échantillonnée (i.e., sa résolution est réduite à la moitié, cf. Figure 2.10a) et est passée au niveau suivant de la pyramide afin de réitérer le processus. Le niveau zéro de la pyramide est l'image en entrée elle-même.

Une représentation multi-échelles par décomposition de pyramide permet de représenter les structures de texture définies à différentes résolutions. Cependant, il y a tendance à perdre l'information à des niveaux très élevés (représentation fine de l'espace d'échelle) où d'un côté la taille de l'image est très petite, et d'autre côté la valeur de l'écart type du filtre Gaussien est assez élevée de sorte que tous les détails fins sont estompés.

La plupart des variantes présentées sont basées sur l'utilisation des filtres de lissage (filtres passe-bas) et suivant différentes conceptions de l'espace d'échelle. Cependant, le principe de représentation multi-échelles demeure classique, i.e., un descripteur LBP est calculé à chaque niveau d'échelle indépendamment des autres niveaux. Le descripteur final est généré par une concaténation de tous les descripteurs locaux LBP par échelle.

Il est important de noter que les travaux réalisés pour la description de la texture d'écorce sont des méthodes initialement appliquées, pour la plupart, sur des textures génériques. Cependant, d'autres méthodes pour la description de la texture existent, notamment, avec d'autres représentations multi-échelles. Ces méthodes ont montré leur efficacité sur diverses bases de textures génériques enregistrant des performances importantes. En particulier, des variantes LBP multi-échelles. En cause de son succès dans plusieurs domaines, plusieurs variantes ont été proposées, où chacune apporte une amélioration à une étape donnée du calcul du code LBP, et la construction du descripteur final.

Par conséquent, dans ce qui suit, on présente une revue des travaux sur l'analyse de texture. Ensuite, une analyse des propriétés de ces approches par rapport aux objectifs de cette thèse, nous permettra d'adapter une méthode pour la description de la texture de l'écorce tout en tenant compte des spécificités de la surface de cet objet naturel.

2.1.1.3 Autres approches de représentation de la texture

Dans le contexte de représentation de la texture, les premiers travaux ont adopté une approche basée essentiellement sur l'apprentissage d'un dictionnaire de textons. Considérés comme des éléments caractéristiques des différentes textures, l'extraction des textons est étroitement liée au développement de nouveaux filtres multi-résolutions et multi-orientations, et l'utilisation des BoW pour générer le descripteur final de la texture. On commence par les approches qui ont proposé des filtres multi-résolutions et multi-orientations pour l'extraction des textons. Par la suite on consacre une partie aux différentes variantes LBP multi-échelles.

2.1.1.3.1 Dictionnaire de textons

Filtres de Leung et Malik (LM)

Les filtres LM de LEUNG et MALIK [2001] sont utilisés pour l'extraction et la description d'attributs locaux pour des images avec des variations de luminance et de prise. Plus particulièrement, LEUNG et MALIK [2001] a introduit une définition des *textons* comme étant les centres des clusters des vecteurs descripteurs d'attributs locaux.

Le principe de cette approche repose sur les images des réponses des filtres obtenues par une opération de convolution de l'image avec un ensemble de filtres. Les filtres de Leung-Malik (LM) LEUNG et MALIK [2001] comportent un ensemble de 48 filtres dont 2 filtres de dérivée Gaussienne selon 6 orientations et 3 échelles, 8 filtres de Laplacien de Gaussienne et 4 filtres Gaussiens (cf. Figure 2.11a).

Filtre de Schmid (S)

Les filtres de Schmid (S) SCHMID [2001] forment un autre ensemble de 13 filtres de Gabor en structure isotopique afin d'assurer une propriété d'invariance à la rotation comme le montre la Figure 2.11b.

Filtre MR8

Les filtres de Varma-Zisserman VARMA et ZISSERMAN [2005], dénotés *Maximum Response 8* (MR8) ont montré des performances supérieures aux deux autres filtres LM et S. MR8 englobe 38 filtres dont 2 filtres anisotropes selon 6 orientations et 3 échelles, et 2 filtres invariants à la rotation dont un filtre Gaussien et un filtre Laplacien de la Gaussienne. Toutefois, 8 filtres réponses seulement sont retenus parmi les 38 selon leurs réponses maximales à chaque échelle à travers toutes les orientations pour les filtres anisotropes (cf. Figure 2.11c).

Ces approches ont permis de représenter des images de textures génériques, prises avec différentes variations photométriques, avec des attributs intrinsèques dérivés de l'espace des réponses des filtres en décrivant les structures caractéristiques comme les contours, les bords, etc. Cependant, la conception d'un ensemble de filtres dépend majoritairement du domaine de l'application (des filtre isotrope ou anisotrope selon le besoin d'une invariance à la rotation ou non, et une analyse multi-échelles et multi-orientations pour une description complète).

Descripteurs de patch

Dans une étude établie par VARMA et ZISSERMAN [2009], les auteurs ont montré que la représentation statistique du modèle de texton directement à partir des valeurs d'intensité des pixels dans un voisinage local, appelé aussi *patch*, produit des résultats comparables voire plus performants que ceux produits par les variantes basées sur les bancs de filtres avec un large support LEUNG et MALIK [2001]; SCHMID [2001]; VARMA et ZISSERMAN [2005].

L'approche proposée, dénommé *Varma-Zisserman-Joint* (VZ-Joint) VARMA et ZISSERMAN [2009], utilise les valeurs d'intensité des pixels dans un voisinage de taille réduite pour construire le dictionnaire des textons.

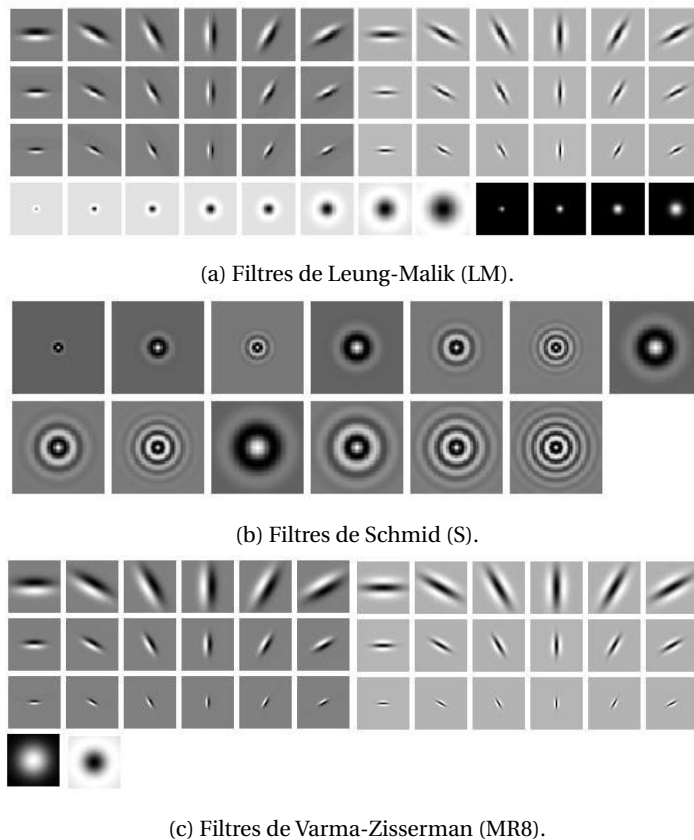


FIGURE 2.11 – Les filtres utilisés pour générer les image réponses pour les construction des dictionnaire des textons VARMA et ZISSERMAN [2005].

2.1.1.3.2 Variantes LBP

Rappelons dans cette partie le principe du calcul d'un code LBP qui passe par trois phases : (i) définition du schéma du voisinage, qui est donné par un support spatial local de 3×3 , (ii) représentation de l'information dans le voisinage, en l'occurrence ici, l'information d'intensité au niveau de gris, et (iii) binarisation et encodage du gradient, qui, pour le LBP est un gradient orienté, où il consiste à prendre le pixel central comme seuil de binarisation des signes des différences d'intensité entre ce dernier et ces voisins autour.

Cette simple représentation et le faible coût de calcul, pour encoder des attributs statistiques de bas niveau, ont fait que le LBP a attiré beaucoup d'intérêt, et par conséquent il était appliqué dans de multiples domaines comme la classification de texture, les systèmes de recherche d'images CBIR MURALA et collab. [2012]; WANG et WANG [2013], la reconnaissance faciale AHO-NEN et collab. [2006]; TAN et TRIGGS [2010], et l'imagerie médicale NANNI et collab. [2010, 2012].

Dans toutes ces applications, les variantes LBP proposées ont démontré une amélioration de performances par rapport aux variantes basées sur les filtres, comme LM, S, MR8 et le VZ-Joint. Spécifiquement, pour les applications de classification de texture sur des bases d'images génériques contenant différentes classes de surfaces naturelles ou synthétiques, le LBP permet une caractérisation robuste et discriminante de l'information de texture LIU et collab. [2017].

Cependant, en dépit de ces propriétés intéressantes, le LBP présente certaines limitations, notamment, pour l'encodage des macrostructures, la sensibilité au bruit, surtout au niveaux d'échelle

élevés, et l'encodage d'autres variations d'information de gradient. Dans la suite, nous exposons certaines variantes LBP en les répertoriant par rapport aux étapes de génération du code LBP. Nous présentons en premier les variantes qui proposent différents schémas de voisinage tout en se concentrant principalement sur un voisinage multi-échelles. Ensuite, nous abordons les différentes représentations de l'information dans le voisinage, suivies des variantes utilisant différents seuils de binarisation et autres encodages du gradient. Enfin, on aborde les variantes LBP basées sur la combinaison de multiples descripteurs locaux pour apporter une information complémentaire. Le Tableau 2.2 présente un récapitulatif de toutes les variantes LBP citées dans cette partie.

Voisinage multi-échelles

Le descripteur LBP est limité à encoder seulement les microstructures. La représentation en multi-résolutions MResLBP a permis de réaliser une représentation des macrostructures, mais à la limite de trois échelles et des motifs *riu2*. Cependant, les motifs *riu2* ne reflètent pas toute l'information de la structure de la texture, et qu'une partie de l'information est comprise dans les motifs non-uniformes attribués tous au même bin. De plus, le MResLBP encode chaque niveau de la résolution indépendamment des autres niveaux ce qui ne permet pas d'encoder l'information inter-échelles, et conduit à une perte de l'information de texture.

Aussi, d'une part, pour encore de larges macrostructures, il va falloir combiner un grand nombre de descripteurs LBP avec des valeurs de voisins, N , élevées ce qui génère un descripteur final d'une très grande dimension. D'autre part, plus la taille du support spatial est grande, plus le risque d'encoder des valeurs bruitées augmente [MÄENPÄÄ et PIETIKÄINEN \[2003\]](#).

Pour remédier à ces limites, des variantes LBP multi-échelles ont été proposées avec la notion de 'patch'. Ces patches forment des régions locales autour du pixel central pour former le nouveau voisinage. Le calcul des patches est le plus souvent réalisé par l'application des filtres passe-bas pour diminuer l'effet de bruit à des niveaux d'échelle élevés.

On retrouve les variantes comme la Three Patch LBP (TPLBP) [WOLF et collab. \[2008\]](#), et la Four Patch LBP (FPLBP) [WOLF et collab. \[2008\]](#) calculent une différence entre les patches voisins représentés par leurs valeurs moyennes. Pixel To Patch (PTP) [WANG et collab. \[2013\]](#) suit le même schéma de représentation de patches mais encode la dérivée seconde entre les patches voisins et le patch central. Dans 'Binary Rotation and Noise Tolerant' (BRINT-SC-MC) [LIU et collab. \[2014\]](#), les auteurs calculent une moyenne dans le sens circonférentiel afin d'assurer un nombre constant de voisins à tous les niveaux de l'espace d'échelle.

[LIU et collab. \[2016b\]](#) proposent une représentation multi-échelles fondée sur l'utilisation de la valeur médiane des pixels contenus dans le patch. Le descripteur Median Robust Extended LBP (MRELBP) montre la stabilité de la médiane pour les régions bruitées de l'image.

Représentation de l'information dans le voisinage

Dans le calcul du descripteur LBP, les valeurs d'intensité (niveau de gris) des pixels du voisinage sont utilisées directement pour le calcul de la différence locale. Cependant, ce processus est relativement sensible aux changements de luminance. Par conséquent, une catégorie de variantes qui encodent une représentation de l'information du voisinage à partir des valeurs d'intensité a

été proposées afin d'améliorer la représentativité de la structure de la texture. Ces variantes génèrent un code LBP soit à partir d'une (i) représentation basée sur la valeur de la différence locale, (ii) une représentation de la magnitude du gradient, ou bien (iii) une représentation statistique de la distribution de l'intensité du voisinage.

Pour la représentation de la magnitude, [GUO et collab. \[2010a\]](#) a décomposé la différence entre les pixels voisins et le pixel central en deux composantes : signe et magnitude. Ces deux informations complémentaires correspondent respectivement aux descripteurs CLBP_S qui est en réalité le LBP original, et le CLBP_M qui encode une information du contraste local du voisinage. La magnitude a aussi été encodée dans de nombreuses variantes LBP comme CLBC-S/M/C [ZHAO et collab. \[2012b\]](#), BRINT-SC-CM [LIU et collab. \[2014\]](#), et le MC-SBP2 [NGUYEN et collab. \[2016\]](#).

Autre que la magnitude de la différence locale, [ZHANG et collab. \[2005\]](#) proposent une représentation de l'image d'intensités par des images de magnitude de Gabor, appelées maps. Ces maps sont obtenues par la convolution de l'image en entrée avec des filtres de Gabor multi-résolutions et multi-orientations, ensuite les codes LBP sont calculés à partir de ces représentations. Le descripteur généré est appelé Local Gabor Binary Pattern (LGBP). La description de la texture avec LBP et les filtres de Gabor est complémentaire. En effet, LBP encode les micro-motifs qui représentent les détails fins de l'image, alors que les filtres de Gabor encodent l'apparence de la texture à travers un large espace d'échelle dans diverses directions.

[VU et CAPLIER \[2012\]](#) a proposé une variante basée sur LBP qui encode l'information de la structure à partir des images de magnitude de gradient. Les images de gradient sont calculées selon différentes directions du gradient dans un voisinage autour de chaque pixel. Ensuite un descripteur LBP par direction est généré pour chaque image de gradient. Cette représentation de l'information à partir de la magnitude et l'orientation du gradient apporte une propriété d'invariance aux changements de luminance, et une représentation directionnelle de l'information de texture. Cependant, cette approche est sensible aux changements de rotation et nécessite un traitement supplémentaire pour assurer une invariance à la rotation dans le cas des applications sur des images naturelles [VU et CAPLIER \[2012\]](#).

Dans [NGUYEN et collab. \[2016\]](#), l'information du voisinage est représentée par des attributs statistiques. Cette représentation approxime la distribution d'intensité dans le voisinage utilisant des moments d'ordre élevés. Les images moments du premier et second ordre sont le résultat d'un calcul des valeurs de moyenne et de variance respectivement dans un voisinage défini. Le descripteur LBP est ensuite calculé à partir de ces images. La représentation conjointe des codes LBP de chaque image moment forme le descripteur Statistical Binary Pattern (SBP2). Une extension de SBP2 vers une représentation multi-échelles et une combinaison avec l'information de la magnitude, donnant le descripteur MC-SBP2.

[NGUYEN et collab. \[2016\]](#) a montré que plus l'ordre des images moment est élevé plus cela contribue à amplifier les variations locales et à introduire plus de bruit. De ce fait, il a argumenté de se restreindre au moment du premier et second ordre comme un compromis entre l'approximation de la distribution et la discrimination du descripteur.

Seuillage et encodage

Lors de l'étape de binarisation et encodage du gradient, la construction d'un code LBP passe par les trois aspects suivants. Le premier, l'utilisation de la valeur d'intensité du pixel central comme seuil de binarisation, et cette valeur n'est pas encodée lors du calcul du code. Le second, une quantification à deux niveaux (quantification binaire) pour la construction de la chaîne binaire. Finalement, l'encodage d'un gradient orienté qui représente seulement la relation entre les pixels voisins et le pixel central.

Les deux premiers aspects causent particulièrement la sensibilité du LBP au bruit, principalement dû aux changements de luminance, tandis que le dernier aspect peut engendrer une perte de l'information de texture car le LBP n'encode pas la relation entre les pixels voisins. Par conséquent de multiples variantes LBP ont été introduites portant des améliorations sur chacun de ces aspects.

Afin de minimiser la sensibilité du LBP au bruit, des variantes proposent l'utilisation de la valeur moyenne du voisinage comme le Improved LBP (ILBP) [JIN et collab. \[2004\]](#), et le Centralized Binary Pattern (CBP) [FU et WEI \[2008\]](#), ou bien la valeur médiane comme dans le Median Binary Pattern (MBP) [HAFIANE et collab. \[2007\]](#) comme seuil de binarisation des différences locales. Ces variantes encodent aussi la valeur du pixel central. Cependant, ceci double la taille finale du descripteur (par exemple, la dimension d'un descripteur $LBP_{R,8}$ est de $2^8 = 256$ bins, et de $2^9 = 512$ bins pour un descripteur $MBP_{R,8}$).

Pour ce qui est de la quantification de la variation d'intensité, [TAN et TRIGGS \[2010\]](#) introduit le Local Ternary Pattern (LTP) avec une quantification ternaire (à trois niveaux) des variations locales afin de remédier aux changements de luminance. [NANNI et collab. \[2010\]](#) propose dans le Local Quinary Pattern (LQP) une quantification à cinq niveaux. Cependant, la détermination des seuils de quantification reste un paramètre primordial à fixer selon un domaine d'application spécifique. Par conséquent, la quantification binaire reste celle la plus souvent utilisée lors de la construction du code binaire dans différentes approches LBP.

D'autres relations de variation entre les pixels voisins appelées différences angulaires ou différences radiales, peuvent encoder une information complémentaire de la texture. La différence angulaire consiste à calculer la variation d'intensité entre une paire de pixels dans un même niveau d'échelle séparés par un angle donné, comme pour les descripteurs Center-Symmetric LBP (CS-LBP) [HEIKKILÄ et collab. \[2006\]](#), Angular différence LBP (ADLBP) [LIU et collab. \[2012a\]](#), Binary Gradient Contours (BGC) [FERNÁNDEZ et collab. \[2011\]](#), transition LBP (tLBP) et directionnel LBP (dLBP) in [TREFNÝ et MATAS \[2010\]](#). Aussi le Local Neighboring Intensity Relationship Pattern (LNIRP) [WANG et collab. \[2013\]](#), et le Circumferential Derivative Pattern (CDP) [WANG et collab. \[2017\]](#) qui encodent la seconde dérivée dans le sens circonférentiel. Ces variantes encodent un gradient circonférentiel pour générer des attributs LBP complémentaires.

Pour la différence radiale, c'est la différence entre des pixels répartis dans des niveaux d'échelle de la macrostructure, selon une direction donnée, on cite le descripteur Radial Différence (RDLBP) [LIU et collab. \[2012a\]](#). RDLBP montre un apport supérieur par rapport à la différence angulaire ADLBP, et est utilisé dans [LIU et collab. \[2016a,b\]](#). [WANG et collab. \[2017\]](#) aussi propose une variante dite Radial Derivative Pattern (RDP) basée sur la seconde dérivée. Le gradient encodé par ces variantes est dit gradient radial.

Une forme plus générique du gradient orienté de LBP est introduite dans ZHANG et collab. [2010], nommée Local Derivative Pattern (LDP) via l'encodage des relations directionnelles entre les pixels voisins basées sur des dérivés locaux d'ordre élevé. Le LDP considère le LBP comme dérivé du premier ordre. Toutefois, le LDP est sensible aux variations de rotation, aussi, le calcul des variantes LDP d'ordres supérieurs permet d'extraire plus de détails mais à un coût de calcul et de représentation important.

Combinaison d'information complémentaire

La combinaison d'attributs LBP issus de différents descripteurs locaux a été proposée en premier dans CLBP_S/M/C GUO et collab. [2010a]. Le descripteur Completed LBP CLBP_S/M/C est une représentation conjointe de trois descripteurs décrivant respectivement le signe CLBP_S, la magnitude CLBP_M de la différence locale, et le pixel central CLBP_C. Le descripteur CLBP_C encode la variation d'intensité du pixel central par rapport à la moyenne globale de l'image. Cette représentation permet d'encoder différents aspects de l'apparence de la texture, à savoir, la variation d'intensité, le contraste local, et la distribution globale du niveau de gris. La combinaison de ses attributs offre une description complémentaire de la texture GUO et collab. [2010a]. Cependant, CLBP_S/M/C GUO et collab. [2010a] a des inconvénients relatifs aux attributs LBP très locaux qui n'encode pas les macrostructures, et sa sensibilité au bruit, et aux changements de rotation, et à sa dimension élevée, en raison de la représentation conjointe.

Par conséquent, plusieurs variantes LBP ont repris la description de la texture suivant le schéma de CLBP_S/M/C GUO et collab. [2010a] avec des améliorations en terme d'invariance au bruit et à la rotation, et la réduction de la dimension du descripteur final.

ZHAO et collab. [2012b] introduise le Completed Local Binary Count CLBC-S/M/C. À la différence de CLBP_S/M/C GUO et collab. [2010a], le CLBC_S/M/C ZHAO et collab. [2012b] n'encode pas de gradient orienté à partir de la chaîne binaire LBP, mais retourne le nombre d'occurrence de '1' dans cette chaîne. Cette stratégie d'encodage assure une invariance à la rotation locale de la structure. De plus, la dimension est considérablement réduite de $2^{2N} \times 2$ à $2(N+1)^2$. Par contre, CLBC_S/M/C ZHAO et collab. [2012b] reste considérablement sensible au bruit.

LIU et collab. [2014] propose BRINT_SC_CM. Afin d'ajouter une invariance à la rotation, BRINT_SC_CM utilise les motifs ri . De plus, sa représentation du voisinage dans l'espace d'échelle permet d'encoder l'information multi-échelles avec une dimension constante à chaque niveau avec une robustesse au bruit grâce à l'utilisation d'un filtre moyen circonférentiel.

Par ailleurs, la combinaison d'attributs LBP, calculés par différents gradients, permet de représenter la texture à travers la description de différentes relations dans le voisinage. Le NI/RD/CI LIU et collab. [2012a] combine trois descripteurs LBP encodant chacun une représentation du contraste local, un gradient radial, et la représentation de la distribution globale du niveau de gris (similaire au CLBP_C GUO et collab. [2010a]). Une extension de NI/RD/CI LIU et collab. [2012a] pour représenter la texture aux niveaux micro et macro était proposée dans MRELBP LIU et collab. [2016b], où la représentation du voisinage est basée sur l'utilisation des patches médians. Dans ELBP LIU et collab. [2016a], un ensemble de six descripteurs binaires qui encodent le signe et la valeur de différence des trois opérateurs suivants : le LBP, le ADLBP avec un gradient circonférentiel, et le RDLBP avec un gradient radial. La concaténation des six descripteurs forme le descripteur final

TABLEAU 2.2 – Résumé des différentes variantes LBP.

Catégorie	Remarques	Nom Complet	Abréviation	Ref
LBP basique	Basique	Local Binary Pattern-carré	$LBP_{R,N} - sqr$	OJALA et collab. [1996]
		Local Binary Pattern-circulaire	$LBP_{R,N} - circ$	OJALA et collab. [2002]
	Mapping	Uniform LBP	$LBP_{R,N}^{u2}$	OJALA et collab. [2002]
		Rotation Invariant LBP	$LBP_{R,N}^{ri}$	OJALA et collab. [2002]
		Rotation invariant Uniform LBP	$LBP_{R,N}^{riu2}$	OJALA et collab. [2002]
Voisinage Multi-échelles	Conventionnel	Multi-résolutions LBP	MResLBP	OJALA et collab. [2002]
	Patches	LBP Filtering	LBPf	MÄENPÄÄ et PIETIKÄINEN [2003]
	Gaussien	Pyramid-based LBP	PLBP	QIAN et collab. [2011]
		Multi Block LBP	MB-LBP	LIAO et collab. [2007]
	Patch Moyen	Three/Four Patch LBP	TPLBP/FPLBP	WOLF et collab. [2008]
		Pixel To Patch	PTP	WANG et collab. [2013]
	Patch Median	Median Robust Extended LBP	MRELBP	LIU et collab. [2016b]
Représentation information	Magnitude	Completed LBP-Magnitude	CLBP-M	GUO et collab. [2010a]
		Local Gabor Binary Pattern	LGBP	ZHANG et collab. [2005]
	Gradient	Local Derivative Pattern	LDP	ZHANG et collab. [2010]
	Distribution	Multi-scale Completed SBP 2	MC-SBP2	NGUYEN et collab. [2016]
Seuillage et encodage du gradient	Seuil	Improved LBP	ILBP	JIN et collab. [2004]
		Centralized Binary Pattern	CBP	FU et WEI [2008]
		Median Binary Pattern	MBP	HAFIANE et collab. [2007]
	Quantification	Local Ternary Pattern	LTP	TAN et TRIGGS [2010]
		Local Quinary Pattern	LQP	NANNI et collab. [2010]
	Gradient circonférentiel	Center-Symmetric LBP	CS-LBP	HEIKKILÄ et collab. [2006]
		Angular différence LBP	ADLBP	LIU et collab. [2012a]
		Binary Gradient Contours	BGC	FERNÁNDEZ et collab. [2011]
		transition/directionnal LBP	tLBP/dLBP	TREFNÝ et MATAS [2010]
		Local Neighboring Intensity Relationship Pattern	LNIRP	WANG et collab. [2013]
		Circoferential Derivative Pattern	CDP	WANG et collab. [2017]
	Gradient radial	Radial Difference LBP	RDLBP	LIU et collab. [2016a,b, 2012a]
		Radial Derivative Pattern	RDP	WANG et collab. [2017]
Information complémentaire	Attributs LBP	Completed LBP	CLBP-S/M/C	GUO et collab. [2010a]
		Completed Local Binary Count	CLBC-S/M/C	ZHAO et collab. [2012b]
		Binary Rotation Invariant and Noise Tolorent	BRINT-SC-CM	LIU et collab. [2014]
		Extended Local Binary Pattern	NI/RD/CI	LIU et collab. [2012a]
		Extended Local Binary Pattern	ELBP	LIU et collab. [2016a]
		Median Robust Extended LBP	MRELBP	LIU et collab. [2016b]
	Autres attributs	Local Binary Pattern/Variance	LBP/Var	OJALA et collab. [2002]
		Local Binary Pattern Variance	LBPV	GUO et collab. [2010b]
		Local Binary Pattern Histogram Fourier	LBP-HF	AHONEN et collab. [2009]
		LBP Histogram Fourier Sign-Magnitude	LBP-HF-S-M	ZHAO et collab. [2012a]
		Adaptive Local Binary Pattern	ALBP	GUO et collab. [2010c]
		LBP and Dominant Neighnborhod Structure	LBP-DNS	KHELLAH [2011]

Extended LBP (ELBP). La représentation multi-échelles de ELBP utilise le principe des moyennes circonférentielles introduit dans BRINT_SC_CM LIU et collab. [2014] pour maintenir une dimen-

sion constante à chaque niveau d'échelle.

Par ailleurs, le descripteur LBP a été combiné avec d'autre descripteur de texture. **OJALA et collab. [2002]** propose LBP/Var avec une représentation conjointe de LBP et de la variance local (Var). La variance encode une information de contraste avec une propriété d'invariance à la rotation locale. Cependant, la quantification de la valeur continue de la variance nécessite une étape d'apprentissage pour déterminer des seuils de quantification. Par la suite, **GUO et collab. [2010b]** propose d'utiliser la valeur de la variance pour régulariser les poids des codes LBP dans l'histogramme final, donnant le descripteur LBP Variance (LBPV).

AHONEN et collab. [2009] proposent le LBP Histogram Fourier (LBP-HF) en combinant le descripteur LBP avec la transformée de Fourier (DFT). Les attributs LBP-HF sont calculés à partir de l'histogramme des motifs uniformes LBP^{u2} afin d'assurer une propriété d'invariance à la rotation globale. Plus tard, le calcul de DFT était combiné avec CLBP_S et CLBP_M **GUO et collab. [2010a]** pour former une représentation complémentaire LBPHF_S_M **ZHAO et collab. [2012a]**.

2.1.2 Représentation de la couleur d'écorce

Les approches de texture présentées jusqu'ici opèrent toutes dans l'espace monochrome (niveaux de gris) pour décrire la texture de l'écorce indépendamment de l'information couleur. Toutefois, la couleur de l'écorce dépend majoritairement des conditions naturelles et de l'environnement de développement de la plante. De ce fait, elle ne constitue pas un attribut pertinent est discriminant en soi en étant très sensible aux changements de luminance, mais, la couleur peut être combinée avec la texture pour apporter une information complémentaire pouvant aider à l'identification de son espèce.

De manière générale, pour encoder l'information de la couleur, on retrouve deux façons de procéder. La première approche consiste à appliquer une méthode de description de la texture sur chacun des plans de l'espace couleur, ou bien uniquement sur un plan en particulier. La description finale est formée par la combinaison de tous les descripteurs de texture par plan. La seconde approche consiste à combiner un descripteur de texture et un descripteur de couleur par une simple concaténation pour formuler la représentation finale de l'image d'écorce.

Commençant par les approches de description de la texture à partir d'une image couleur, où il est essentiel de déterminer l'espace de couleur qui permet de différencier les écorces à partir de leurs structures colorées. Toutefois, étant l'espace des couleurs primitives, l'espace RGB est un espace référentiel dans la plupart des travaux qui évaluent l'ajout de l'information de la couleur, comme **WAN et collab. [2004]**, qui a comparé les performances des GLCM extraites en niveaux de gris avec celles extraites des images RGB d'écorce. Les auteurs ont obtenu une amélioration significative des performances. **BRESSANE et collab. [2016, 2015b]** ont encodé la texture de l'écorce dans le plan V de l'espace HSV utilisant les GLCM et des statistiques calculées à partir de l'histogramme de l'image.

Dans **HUANG et collab. [2006c,d]** la texture est représentée par les filtres de Gabor et les Contourlets (CT) pour différents espaces de couleur et comparée avec le niveau de gris. Les meilleurs scores de classification ont été obtenus avec l'espace YIQ. Dans **HUANG [2006]**, l'énergie des coefficients issus d'une décomposition en DWT dans l'espace YCbCr et en niveau de gris sont combinés. Les résultats retrouvés ont montré l'apport de l'information de la couleur pour améliorer les scores

de classification.

Une représentation dense à base de points d'intérêt SIFT dans différents espaces de représentation de la couleur était utilisée par NLab UTokyo [NAKAYAMA \[2013\]](#) avec des variantes SIFT, à savoir, C-SIFT, Opponent-SIFT, HSV-SIFT dans PlantNet 2014.

Pour ce qui est de la deuxième façon d'incorporer l'information couleur dans la représentation de la surface d'écorce, la description couleur la plus évaluée est celle de la concaténation des histogrammes des plans d'un espace donnée. [BERTRAND et collab. \[2017\]](#) a représenté la couleur de l'écorce par l'histogramme du plan H de l'espace HSV. Le plan H permet de représenter les écorces avec des nuances jaunâtres sur 256 valeurs [BERTRAND et collab. \[2017\]](#). Cet histogramme est combiné avec les descripteurs de texture pour former la description finale de l'écorce.

PlantNet [BAKIĆ et collab. \[2013\]](#) a utilisé une combinaison d'histogrammes couleur : RGB pondéré $wRGB$ et l'histogramme de HSV dans ImageCLEF 2013, et le $wRGB$ et $wLUV$ dans PlantCLEF 2014 [GOËAU et collab. \[2014b\]](#).

Autres descripteurs couleur ont été aussi utilisés pour caractériser l'information de la couleur comme la distribution des angles (*Color Distribution Angles* (CDA)) dans l'approche de [HUANG et collab. \[2006b\]](#), et un descripteur couleur basé sur les moments (Color Moments) évalué par IBM AU [CHEN et collab. \[2014\]](#) lors de PlantCLEF 2014.

2.1.3 Classification des écorces

Dans un système d'identification de plantes, la détermination de l'espèce d'une plante consiste à mettre en correspondance une image requête avec des images références de la base. Pour ce faire, une comparaison est établie entre le descripteur de l'image requête et ceux des images de références. Cette mise en correspondance peut suivre deux schémas possibles, à savoir, une recherche du contenu visuel par similarité, comme le cas des systèmes CBIR (*Content-Based Image Retrieval*), ou bien une classification via une méthode d'apprentissage supervisé.

Dans un système de recherche par contenu visuel, le processus de recherche par similarité consiste à utiliser une base d'images étiquetées en sa totalité pour déterminer l'espèce d'une image requête, jusque là inconnue. La comparaison de tous les descripteurs des images de la base individuellement avec le descripteur de l'image requête permet de retourner une liste des images les plus similaires selon la notion du voisin le plus proche (*Nearest Neighbor*). L'image requête est associée alors à la même espèce de l'image réponse la plus similaire [CHI et collab. \[2003\]](#); [GOËAU et collab. \[2012\]](#); [KIM et collab. \[2011\]](#); [TSCHIEDEL GUEDES et BORGES \[2018\]](#).

La mesure de similarité entre les images est estimée par le calcul d'une distance entre leurs descripteurs. Le choix du descripteur ainsi que la distance influence directement le processus de recherche. Selon l'échelle de représentation de l'image, globale ou locale, un calcul de similarité est approprié. En effet, une représentation globale de l'image est caractérisée par un seul vecteur descripteur. Le calcul de similarité entre les descripteurs de la base et celui de l'image requête est ainsi global. Cependant, ce calcul peut s'avérer coûteux lorsque la dimension du vecteur descripteur est grande. Toutefois, des solutions de réduction d'espace comme l'analyse par composante principale (*Principal Component Analysis PCA*) [CHEN et collab. \[2014\]](#) ou analyse par corrélation canonique (*Canonical Correlation Analysis CCA*) [NAKAYAMA \[2013\]](#) préalablement au calcul de similarité peuvent réduire considérablement le processus de recherche.

Dans le cas d'une représentation locale, l'image est représentée par un ensemble de descripteur locaux (cas des descripteurs de points d'intérêt). Le calcul de similarité revient donc à faire un appariement (*matching*) entre des paires de descripteurs requête/réponse. De même pour le calcul de similarité globale, des techniques d'optimisation du temps de calcul peuvent réduire le temps de recherche comme les tables de hachage (*Hashing tables*) BAKIĆ et collab. [2013]; GOËAU et collab. [2014b].

Dans le cas d'une représentation globale des descripteurs locaux par le modèle du sac des mots (BoW) WENDEL et collab. [2011], le calcul de similarité est ainsi effectué comme dans le premier cas de la représentation globale.

Procéder à une recherche par similarité pour déterminer l'espèce d'une image d'écorce modélise bien le cas de pratique réelle. En effet, pour déterminer l'espèce d'une image, les naturalistes procèdent à une recherche par similarité parmi les espèces connues, pour trouver celle qui partage des caractéristiques assez proches que celles de l'espèce recherchée. Ainsi, à partir de ces espèces similaires retrouvées, ils décident de l'appartenance de cette plante à telle ou telle espèce.

Lorsque l'identification de l'espèce passe par une classification, des méthodes d'apprentissage automatique sont utilisées. Le principe de classification consiste à prédire l'espèce d'une image à travers un système entraîné sur la base des descripteurs des images d'un sous ensemble de la base dit ensemble d'apprentissage (*training set*). L'image requête est dans ce cas un exemple non appris auparavant par le système, et faisant partie d'un sous ensemble de la base d'images dit ensemble de test (*testing set*).

Les méthodes utilisées dans le contexte de classification des écorces sont les k plus proches voisins (k -NN) BRESSANE et collab. [2015b]; GOËAU et collab. [2014b]; SIXTA [2011]; SONG et collab. [2004]; WAN et collab. [2004], les réseaux de neurones [FEKRI-ERSHAD, 2020], et les réseaux à base radial probabiliste (*Radial Base Probabilistic Neural Network (RBPNN)*) HUANG [2006]; HUANG et collab. [2006b,c,d,d], les machines à support de vecteurs (*Support Vector Machine(SVM)*) BAKIĆ et collab. [2013]; BERTRAND et collab. [2017]; BLANCO et collab. [2016]; CHEN et collab. [2014]; LE et collab. [2015]; ŠULC et MATAS [2013]; SULC et MATAS [2014]; ŠULC et MATAS [2017]; WENDEL et collab. [2011]; ZHAO et collab. [2015]; ZHU et collab. [2017], la régression logistique (*Linear Logistic Regression*) NAKAYAMA [2013], les forêts aléatoires (*Random Forest (RF)*) OTHMANI et collab. [2013], et un modèle statistique inspiré des RF pour construire un arbre de décision binaire dans BRESSANE et collab. [2015a, 2016].

Les performances d'une méthode d'apprentissage sont étroitement liées à la qualité des descripteurs utilisés pour caractériser le contenu visuel des images.

2.2 Classification par apprentissage profond

L'apprentissage profond (*Deep Learning (DL)*) est une forme avancée de l'apprentissage automatique à base de réseaux de neurones. Il est constitué de plusieurs couches de traitement capables d'apprendre des caractéristiques et attributs représentatifs du contenu de façon automatique, sans pour autant faire la conception de la primitive à encoder, et de tout le processus de son extraction, encodage et représentation. Ces attributs sont appris à partir d'un très grand nombre d'échantillons de données présentés au système d'apprentissage profond.

Selon la classe de l'apprentissage, i.e., supervisé, non-supervisé, par renforcement, plusieurs modèles d'apprentissage profond sont définis. Dans le contexte de notre étude, nous nous intéressons à la classe supervisée. Parmi les modèles DL adaptés en grande partie pour la reconnaissance des images, on trouve les réseaux de neurones à convolution (*Convolutional Neural Network*(CNN)).

Dans la suite de cette section, nous donnons une brève présentation des CNN, leur entraînement et utilisation du principe de transfert d'apprentissage, au final, nous citons les travaux à base de CNN dans le contexte de la classification des écorces.

2.2.1 Réseaux de neurones à convolution CNN

Les réseaux à convolution sont un type particulier des réseaux de neurones. Ils combinent une architecture particulière et un entraînement de bout en bout par la rétro-propagation du gradient.

Le processus d'apprentissage est essentiellement basé sur l'optimisation d'une fonction de coût pour réduire la valeur d'erreur entre les classes prédites et les classes réelles de la vérité terrain. L'adaptation des paramètres suit le principe de la rétro-propagation de gradient [BENGIO et collab. \[2013\]](#); [GOODFELLOW et collab. \[2016\]](#); [LECUN et collab. \[2015\]](#). Ainsi, c'est l'architecture du réseaux qui change. Un réseaux convolutif typique est constitué d'un empilement de trois types de couches : convolution, fonction non-linéaire d'activation, et une couche d'agrégation (pooling).

La couche de convolution *Conv* consiste en l'application de l'opération de convolution entre un certain nombre de filtres linéaires et des cartes de caractéristiques (les premières cartes sont les trois plans RGB de l'image en entrée). Ensuite une application d'une fonction non linéaire, comme *ReLU* ou *Sigmoid* qui va générer une carte d'activation. Au final, une couche de pooling (*Pool*), dont l'intérêt est de compacter d'avantage la dimension des cartes de caractéristiques, et apporter une invariance à la translation est appliquée. La couche supérieure d'un CNN est un simple réseau de neurones multi-couches qui fait office de classifieur afin de déterminer une probabilité d'appartenance à une classe donnée de la base d'images utilisée. Cette dernière couche est appelée une couche complètement connectée *Fully connected* (*FC*). Ainsi, un réseau convolutif à l'architecture typique suivante : $Conv \rightarrow ReLu \rightarrow Pool$.

Ces trois couches peuvent être répétées un certain nombre de fois. Plus on a de couches plus le réseau est profond et plus le nombre de paramètres à apprendre accroît. Par conséquent, plusieurs modèles de CNN, à chacun une architecture particulière, sont introduits, et cela, suite aux remarquables résultats de classification en 2012 de AlexNet [KRIZHEVSKY et collab. \[2012\]](#) sur le large corpus d'images *ImageNet* avec 1000 classes différentes et plus de 1.2 million d'images d'apprentissage [RUSSAKOVSKY et collab. \[2015\]](#).

Parmi les modèles CNN les plus utilisés, on cite AlexNet [KRIZHEVSKY et collab. \[2012\]](#), VGGNet [SIMONYAN et ZISSERMAN \[2014\]](#), GoogleNet [SZEGEDY et collab. \[2015\]](#), ResNet [HE et collab. \[2016\]](#), et DenseNet [HUANG et collab. \[2017\]](#). AlexNet étant le plus ancien, et le plus simple. En effet, ces modèle vont du plus simple au plus profond et au plus complexe.

Les CNN ont de nombreuses applications comme la reconnaissance, la détection, la localisation d'objet, et bien d'autres applications [GOODFELLOW et collab. \[2016\]](#).

2.2.2 Utilisation des CNN

Les CNN sont conçus pour apprendre un très grand nombre de paramètres sur des bases volumineuses d'images annotées, où les méthodes classiques atteignent leurs limites. Toutefois, ceci pose une contrainte majeure quant à l'applicabilité et l'efficacité des CNN, qui requièrent un très grand nombre d'images pour apprendre le modèle.

Cependant, parmi les avancées les plus notables dans le domaine, qui permettent de dépasser cette limite de collecter de grandes bases annotées, est le transfert d'apprentissage (*transfer learning*). En effet, un CNN entraîné sur une très grande base d'images peut *transférer* ses paramètres appris vers d'autres problèmes sur lesquels il n'était pas entraîné au préalable, et ceux avec un moindre coup d'adaptation des paramètres. PAN et YANG [2009] présente dans sa revue tous les détails du transfert d'apprentissage avec toutes les techniques sous-jacentes.

C'est dans cette perspective que la plus part des auteurs des modèles CNN choisissent de partager les paramètres appris sur de grandes bases, comme *ImageNet*, dans le but que ces modèles seront ré-utilisés pour extraire des attributs ou bien seront réadaptés sur d'autres problèmes, où il n'y a pas suffisamment d'images pour entraîner efficacement ce modèle à partir de poids aléatoires.

Le premier cas d'usage d'un CNN, pré-entraîné comme extracteur d'attributs, est généralement adapté pour des bases de faible taille, mais similaires à la base sur laquelle le réseau convolutif était entraîné. Il s'agit ici de former le vecteur descripteur d'une image donnée. Formellement, un CNN est composé de couches, la sortie de chaque couche est un ensemble de cartes de caractéristiques qui forment un vecteur descripteur. Ainsi pour représenter le contenu visuel d'une image, on la passe dans le CNN à travers les différentes couches, où des séries de convolution, activation, pooling sont appliquées. Généralement, le vecteur descripteur de l'image est extrait à partir de la dernière couche FC. Cette couche à une représentation globale et abstraite du contenu visuel, à l'encontre des couches inférieures qui encodent des attributs locaux de bas niveau. Cependant, cette représentation globale peut ne pas se révéler très pertinente si l'image est très différentes des images d'apprentissage du CNN CIMPOI et collab. [2016]. La classification de cette image est par la suite réalisée par un classifieur SVM linéaire afin de déterminer sa classe.

Le second cas d'usage d'un CNN pré-entraîné avec réadaptation de ses poids est adapté pour des bases de faible taille, mais très différentes de la base sur laquelle le CNN était entraîné au préalable. La représentation de l'image est réalisée de bout-en-bout en transférant l'apprentissage du CNN pour une ou plusieurs couches intermédiaires.

Lorsqu'on procède ainsi, la dernière couche, FC, est généralement modifiée de sorte à avoir un nombre de nœuds qui correspond au nombre de classes de la nouvelle base sur laquelle on veut appliquer le CNN. Pour les couches intermédiaires, il est de pratique de *geler* les couches basales. En effet, ces couches encodent des attributs de bas niveaux qu'on peut généraliser sur différents contenus visuels (des contours, des coins, etc). Ainsi, geler ses couches revient à transférer les valeurs de leurs poids vers la nouvelle application du CNN. Ainsi, le processus d'apprentissage dans ce cas consiste à apprendre les nouveaux poids pour la nouvelle couche FC et à réadapter les poids pré-appris de certaines couches intermédiaires pour les images de la nouvelle base.

2.2.3 Classification des écorces par CNN

L'utilisation des CNN pour la classification des écorces était, jusque là, limitée par la disponibilité d'une base publique d'écorce d'une grande taille. Récemment, [CARPENTIER et collab. \[2018\]](#) a introduit une nouvelle base des espèces d'écorce avec plus de 20000 images, et a proposé l'utilisation des réseaux de neurones à convolution pour la classification des espèces d'écorce.

Le modèle résiduel ResNet34, pré-entraîné sur la base *ImageNet*, a été utilisé tout en réadaptant ces poids sur la nouvelle base d'écorce. En effet, avec des bases d'images de taille relativement inférieure à celle de la base *ImageNet*, il est préférable de faire un transfert d'apprentissage. Les résultats obtenus par [CARPENTIER et collab. \[2018\]](#) montrent l'intérêt d'un apprentissage automatique des attributs pertinents pour la description de la texture d'écorce, notamment pour distinguer les différentes espèces avec une forte similarité inter-classes.

Néanmoins, le choix des auteurs du modèle CNN, le ResNet34, n'est pas issu d'une comparaison quantitative entre différents modèles de CNN, car à priori ils n'ont inclus que des variantes du modèle résiduel avec différentes profondeurs, ResNet18 et ResNet34.

2.3 Synthèse et discussion

À partir de cette revue des travaux sur la classification des écorces, nous avons pu tirer les remarques suivantes portant sur l'apport et l'inconvénient de chaque aspect des approches d'identification d'écorce :

- ▶ L'utilisation de l'écorce dans un contexte d'identification multi-organes est une tâche proche du contexte d'application réelle. Cependant, c'est une tâche très contraignante dans un contexte automatisé, notamment, lors de la construction de la base d'images, et la prise de décision sur l'espèce végétale. Cette dernière est compliquée lorsqu'il s'agit de décision à partir de la combinaison de différentes sources d'information sur les organes considérés, et de multiples descripteurs encodant des primitives différentes.
- ▶ Une vaste majorité des approches de description de l'écorce reposent sur l'utilisation des méthodes génériques de la texture et de la couleur, et peu de travaux ont approché le problème par une modélisation d'une description botanique, et ceux du fait que l'apparence de la surface de l'écorce est très variable en texture et en couleur ce qui rend l'extraction d'attributs spécifiques une tâche difficile.
- ▶ Le caractère visuel essentiellement utilisé pour la description de la surface de l'écorce dans tous les travaux est la texture. Compte tenu de la similarité inter-espèces, cela montre que la texture est porteuse d'information et permet de distinguer entre les différentes espèces végétales à l'encontre de la couleur.
- ▶ On distingue deux principales représentations des textures d'écorce : une représentation des structures locales, et une représentation multi-échelles à partir de larges supports spatiaux, en utilisant une analyse multi-résolutions et/ou multi-orientations. Cette dernière englobe le plus de travaux, et montre un intérêt à encoder l'information macrostructure pour la description des images d'écorce.

- ▶ La représentation multi-échelles à base de points d'intérêt s'avère concurrente est assez prometteuse surtout avec un modèle global BoW, ou bien une représentation dense qui capte tous les détails de l'information de texture, mais assez coûteuse en temps de calcul. Toutefois, cette approche peut conduire à l'extraction et la description de points d'intérêt bruités dans le cas des surfaces abîmées, par des blessures, ou même la présence de mousse ou lichens.
- ▶ La représentation multi-échelles par des méthodes à base de filtres spatio-fréquentiels, comme des variantes d'ondelettes, sont assez utilisées pour la description des objets naturels. Cependant, la conception et le coût de calcul de l'analyse multi-échelles et multi-orientations laisse favoriser les méthodes structurales, basées sur les motifs locaux LBP, qui encodent directement l'information d'intensité de texture.
- ▶ L'application de différentes variantes LBP multi-échelles, montre l'intérêt global de cette approche mais aussi son efficacité en terme de simplicité de modélisation, et sa généralisation pour la description du contenu visuel naturel des écorces. En effet, les variantes LBP multi-échelles ont donné des résultats très intéressants, et ont dépassé les GLCMs et les méthodes basées sur les ondelettes, et aussi les approches à base de points d'intérêt. Cependant, des modifications sont nécessaires pour les adapter à la description des motifs relatifs à la surface de l'écorce, à savoir, des macrostructures.
- ▶ Concernant d'autres approches de représentation de la texture, on a constaté un intérêt majeur à l'approche LBP de la part de la communauté de l'analyse de texture, d'où le développement de plusieurs variantes, notamment, des variantes multi-échelles afin de palier la limite du descripteur basique avec son support très local. Cependant, il est important de noter que pour la plupart de ces variantes, l'espace d'échelle était limité à trois niveaux, avec un mapping *riu2* et huit points voisins par niveau.

Ces choix semblent justifier une contrainte de dimensionnalité du descripteur final. En effet, pour cet exemple de conception, le descripteur final aura une dimension de $(8 + 2) * 3 = 30$, or un mapping *riu2* n'est pas forcément suffisant pour représenter l'information de texture, car il condense tous les codes *nonriu2* dans un seul bin de l'histogramme final. Cependant, garder la représentation originale avec 2^N ($N = 8$ dans *cecas*) par niveau d'échelle donnera un descripteur de $2^8 * 3 = 768$, cette taille augmente d'avantage avec plus d'échelles (par exemple pour 10 échelles, on obtient une dimension de 2560). De plus, on mentionne le nombre de points voisins fixé à huit voisins pour tous les niveaux d'échelle, là encore, et sachant que le LBP encode un gradient orienté dont l'encodage dépend fortement du nombre de points dans le voisinage, appliquer la représentation de [OJALA et collab. \[2002\]](#) avec $N_i = 8 \times R_i$ pour le i^{eme} niveau d'échelle donnera un descripteur final de la dimension très considérable de $2^8 + 2^16 + 2^24 = 16.843.008$. Cette dimension se voit augmentée encore avec plus d'échelles. Ajoutons à cela le risque d'encoder du bruit avec un nombre élevé de points voisins pour un grand support spatial, d'où l'utilisation des filtres locaux (moyen, médian, gaussien, etc), sauf que l'application de ces derniers est fortement liée à l'application.

Les efforts d'encoder d'autres informations dans le support spatial, comme la magnitude, le contraste global, le gradient radial combiné au gradient orienté, soit conjointement ou par

concaténation, améliorent d'avantage le descripteur LBP, mais leur utilisation pour une représentation multi-échelles est confrontée aux mêmes contraintes décrites précédemment. Toutes ces contraintes laissent penser à une nouvelle conception de l'espace d'échelle, une représentation de l'information d'intensité, et un encodage qui permet à la fois de représenter des attributs de texture pertinents et générer un descripteur de macrostructure, dont la taille est indépendante du nombre de points voisins.

- ▶ L'exploitation de l'information couleur est le plus souvent réalisée par l'encodage d'attributs de texture par plan dans un espace de couleur donné. Cette approche a l'avantage de représenter différents aspects de la structure de l'écorce, mais triple la dimension du descripteur final de l'image d'écorce.
- ▶ La description de la couleur dans les espaces de couleur luminance-chrominance est plus pertinente que la description dans l'espace des couleurs primaires RGB. Ceci est lié au fait que les plans corrélés de l'espace RGB sont plus sensibles aux changements de luminance par rapport aux autres espaces de couleur comme le HSV, où l'information de chrominance est séparée de celle de la luminance.
- ▶ La plupart des systèmes d'identification sont basés sur des méthodes d'apprentissage supervisé pour classer les espèces d'écorce. Les classifieurs SVM et réseaux de neurones sont les plus utilisés. Tout de même, l'approche la plus intuitive pour un système d'identification de plantes est de procéder par calcul de similarité entre la description d'une image requête, dont on souhaite identifier l'espèce végétale, et l'ensemble des descripteurs des images de la base. L'espèce de l'image requête correspond ainsi à celle de l'image la plus similaire parmi les images de la base retournées.
- ▶ L'expansion des méthodes de classification à base de modèle CNN avec des techniques de transfert d'apprentissage rend leur application intéressante pour la description des images naturelles. En effet, les CNN adoptent un mécanisme d'apprentissage automatique des caractéristiques visuelles, sans pour autant qu'elles soient modélisées au préalable. Cependant, l'utilisation du transfert d'apprentissage pour les CNN à partir d'une base complètement différente des bases d'écorce ne permet pas une évaluation compréhensible de cet outil. Il convient donc d'entraîner un modèle CNN sur des bases d'écorce afin de pouvoir examiner ses performances, et éventuellement, pouvoir comparer les deux approches de représentation et de classification du contenu visuel, i.e., approche classique avec les deux phases, et approche automatique de bout-en-bout.

2.4 Contexte du travail

Un cas optimal d'identification et classification des espèces végétales est de fournir des images d'organes en entrée à un système automatique d'identification, et ce dernier se charge de retourner leurs espèces potentielles. Cependant, un tel procédé reste complexe et difficile à mettre en œuvre, et est par conséquent toujours en cours de recherche.

Les systèmes existants d'aide à l'identification automatique des plantes se distinguent non seulement par l'organe végétal utilisé, mais aussi, pour cet organe, certaines restrictions sur les

images traitées sont imposées. Dans ce contexte, et parmi les systèmes existants, on trouve la compagnie d'évaluation Pl@antNet [GOËAU et collab. \[2015\]](#) qui regroupe des travaux d'identification à partir de plusieurs organes. Aussi, l'exemple de l'application mobile LeafSnap [KUMAR et collab. \[2012\]](#) qui ne traite que des images de feuilles acquises sur un fond uniforme blanc, l'application Rêves [CERUTTI et collab. \[2013\]](#) qui propose une identification dans le milieu naturel guidée par une segmentation semi-automatique, et au final, [MZOUGHJI et collab. \[2016\]](#), qui réalise une identification spécifique basée sur des connaissances botaniques sur des feuilles prises uniquement sur fond uni.

Pour la plupart des systèmes d'identification à partir des images d'écorce, l'image en entrée ne contient que le tronc de l'arbre, car, à l'encontre des autres organes, il est relativement facile de prendre une image d'une région de la surface du tronc d'arbre sans inclure un fond naturel. D'autres systèmes n'imposent pas des restrictions lors de la prise d'images, et appliquent des traitements à priori comme un découpage automatique de bord [HUANG et collab. \[2006c\]](#), un principe de fenêtrage pour déterminer les régions d'intérêt [LE et collab. \[2015\]](#), ou bien une segmentation automatique afin de réduire l'effet de bruit inclus par le fond naturel [CHEN et collab. \[2014\]](#).

L'application proposée dans ce travail conserve le cadre de prise d'image d'écorce dans son milieu naturel. Toutefois, certaines restrictions sont imposées sur le contenu et la qualité de l'image d'écorce. En effet, on cherche à exploiter les variations des structures présentes sur la surface d'écorce comme élément pouvant les distinguer, et donc le succès de notre approche dépend de la visibilité de ces structures. Toutefois, les restrictions proposées restent parfaitement réalisables par un utilisateur novice dans la cas d'une application réelle :

- **Type d'écorce :** rappelons que la définition de l'écorce est assez générique dans un cadre botanique. Elle se présente comme la couche supérieure protectrice qui enveloppe la tige principale pour les arbustes, le tronc pour les arbres, et les branches. Cependant, l'apparence visuelle de l'écorce est considérablement variable pour ces trois parties (c.f Figure 2.12a), et la possibilité de proposer une solution générique pour leurs descriptions ne semble pas réalisable. Dans notre travail, on désigne par écorce, la couche supérieure qui couvre le *tronc* d'un arbre à un âge mature. Ainsi dit, on exclut de ce traitement, les tiges secondaires, la tige principale pour les arbustes et les très jeune plantes, et les branches secondaires car les caractéristiques visuelles sont largement divergentes, et il ne convient pas d'appliquer la même analyse de contenu.
- **Contenu de l'image :** d'une manière générale, l'image doit contenir essentiellement un tronc d'arbre. Ceci nous évite une étape de prétraitement, i.e., une segmentation complexe et subjective comme pour le cas des exemples illustrés dans la Figure 2.12b). D'un côté, on ne dispose pas d'une vérité terrain pour valider la qualité de la segmentation. De l'autre côté, cette opération peut s'avérer coûteuse, et pratiquement évitable en proposant une distance de prise suffisante pour ne prendre en photo qu'une région d'écorce.
- **Qualité de l'image :** par la qualité des images d'écorce, on fait référence aux problèmes naturels comme les blessures, la présence de la mousse et des lichens qui altèrent potentiellement l'apparence de la surface d'écorce et donc sa structure (c.f Figure 2.12c). Il est important que la structure de l'écorce soit nette pour ne pas inclure une information insignifiante lors de l'analyse du contenu des images d'écorce.

- **Variation d'âge :** un autre problème naturel important à prendre en considération est l'âge de l'arbre et ses effets sur l'apparence visuelle de la surface de l'écorce (c.f Figure 2.12d). Toutefois, il y a une certaine stabilité de l'apparence de l'écorce d'un arbre lorsque il atteint un âge mature. Cette stabilité d'apparence peut s'étaler sur une large période de temps, conservant ainsi les caractéristiques visuelles de la structure d'écorce [GODET \[2012\]](#); [VAUCHER \[1993\]](#); [WOJTECH \[2011\]](#).

En outre, disposer éventuellement d'une information sur l'âge de l'arbre offre la possibilité d'étendre notre étude pour l'identification et de l'espèce d'écorce et de sa tranche d'âge afin d'exprimer l'efficacité d'un tel système face à ce problème naturel bien présent.

Le caractère visuel qui encode l'apparence visuelle de la surface de l'écorce est à priori la texture. L'approche utilisée pour la description de la texture des écorces est une approche structurale et statistique basée sur des motifs binaires.

Le choix des motifs binaires est justifié d'un côté par les performances significatives de cette approche de description de texture dans des applications spécifiques de classification d'écorce [ŠULC et MATAS \[2013\]](#); [SULC et MATAS \[2014\]](#); [ŠULC et MATAS \[2017\]](#) comparée aux autres approches de description de texture, mais aussi dans des applications génériques de classification de texture [LIU et collab. \[2017\]](#) dépassant ainsi en performances toutes les approches basées sur l'apprentissage de textures comme le VS-MR8 [VARMA et ZISSERMAN \[2005\]](#) et le VZ-Joint [VARMA et ZISSERMAN \[2009\]](#).

Le lien entre le choix des motifs binaires et la description botanique est fortement présent sachant que cette dernière utilise la notion de *la structure* d'écorce pour la description de son type botanique, et qu'un descripteur à base de motifs binaires permet une description *structurale* et *statistique* de la texture. Aussi, ce modèle s'adapte dans le cas des textures avec des structures irrégulières en taille et en distribution comme le cas des structures d'écorce.

Le résultat de la caractérisation du contenu visuel est une représentation de l'image de l'écorce par un descripteur. Ce descripteur est une représentation globale des attributs extraits de l'image et sera utilisé pour déterminer son espèce.

La classification de l'espèce végétale d'une image d'écorce est réalisée par un processus de recherche par similarité. Ainsi le problème est approché par une recherche des k plus proche voisin $k - NN$ ($1 - NN$) en comparant les descripteurs des images de la base avec celle d'une image requête. Plus précisément, nous utilisons une stratégie de validation dite *Leave One Out* (LOO), son principe consiste à présenter chaque image dans une base d'écorce comme requête est retourner l'image la plus similaire.

De plus, différentes méthodes d'apprentissage supervisées seront évaluées, à savoir, le $k - NN$ ($1 - NN$), le SVM_{lin} linéaire et le SVM_{rbf} avec un noyau à base radiale. Le choix de ces classifieurs s'inscrit dans le cadre du protocole de comparaison avec les méthodes de l'état de l'art.

Pour ce qui est des méthodes à base de CNN, il convient d'établir une comparaison quantitative entre les méthodes dites *handcrafted*, proposées au cours de cette thèse avec des descripteurs d'attributs profonds extraits de modèles CNN, comme le ResNet34. Cette comparaison permet entre autre de quantifier l'ordre de grandeur des deux approches quant à la classification des images d'écorce, mais aussi d'avoir une idée des limites des descripteurs et classifieurs *handcrafted* sur de grandes bases, soit en terme du nombre d'images, ou en nombre d'espèces d'écorce.



(a) Type d'écorce.



(b) Contenu de l'image d'écorce.



(c) Qualité de l'image d'écorce.



(d) Variation d'âge de l'arbre.

FIGURE 2.12 – Contraintes qui limitent l'application à l'identification et la classification d'écorce.

2.5 Conclusion

Cette partie constitue une exploration de l'existant en terme d'approche dédiées à l'identification et la classification des écorces. Cette revue était suivie d'une revue des approches de classification de textures génériques.

À travers cette revue, on a pu constater les points forts et les points faibles des différentes approches, et avons pu fonder des arguments solides pour défendre notre choix d'une approche pour la description de la texture d'écorce par une approche structurale à base de motifs binaires.

On s'inspire principalement de l'approche LBP. Cette approche a l'avantage d'être facile à étendre et adapter pour un contexte spécifique. Dans notre cas, on propose une approche capable de générer une description des macrostructures caractérisantes pour les différents types d'écorce. L'énoncé de notre approche et sa conception théorique seront exposés en détail dans les parties suivantes de ce manuscrit.

Chapitre 3

Modèle générique pour la représentation de la texture

Sommaire

3.1	Modèle de description structurale	51
3.2	Extraction de la macrostructure	52
3.2.1	Distribution des niveaux d'échelle	52
3.2.2	Distribution des points voisins	53
3.3	Représentation de la macrostructure	55
3.4	Encodage de la macrostructure	57
3.5	Descripteur de la texture d'écorce	59
3.5.1	Représentation globale	60
3.5.2	Représentation par région	61
3.5.3	Représentation locale	61
3.6	Conclusion	62

La définition d'un modèle générique capable de prendre en considération l'importante variabilité des structures d'écorce reste difficile et loin d'être réalisable. Cependant, manipuler une description botanique pour la définition d'un tel modèle montre un intérêt considérable pour mettre en œuvre un outil automatique d'aide à l'identification des espèces d'écorce.

La conception d'une solution pour la description des surfaces d'objets naturels reste une tâche défiante. En effet, la surface des écorces est principalement formée de structures de taille grande est variable, qu'on appelle des *macrostructures*. La nature de ces macrostructures requiert la proposition d'une solution qui opère sur des supports spatiaux assez larges.

Afin de mieux visualiser cette notion de macrostructure, un exemple motivant est présenté dans la Figure 3.1. Il est clair que l'utilisation d'un support spatial large est plus intéressante que celle d'un support très local. Ceci permet de garder l'essentiel de l'information structurale à extraire pour caractériser les différentes structures d'écorce.

Selon cette perspective, on cherche à simuler et modéliser une description botanique pour arriver à une caractérisation efficace et discriminante des macrostructures des différents types d'écorce, en vue de la similarité inter-espèces, et des différents problèmes d'imagerie. Nous proposons une approche spécifique qui adapte une méthode structurale pour la représentation de texture afin d'aboutir à une caractérisation fiable et efficace de la macrostructure de l'écorce. Toutefois, l'efficacité de l'approche proposée repose sur certaines propriétés d'invariance aux conditions d'imagerie.

Le modèle est conçu à l'aide d'outils mathématiques qui permettent une représentation compacte des macrostructures de l'écorce, mais suffisamment porteuse d'information structurale robuste et discriminante pour une application dans des conditions naturelles réelles.

Dans la suite de ce chapitre, on expose de façon générique la définition théorique de notre approche qui consiste essentiellement en *l'extraction*, *la représentation*, et *l'encodage* de l'information de texture pour représenter le motif caractéristique de la macrostructure.

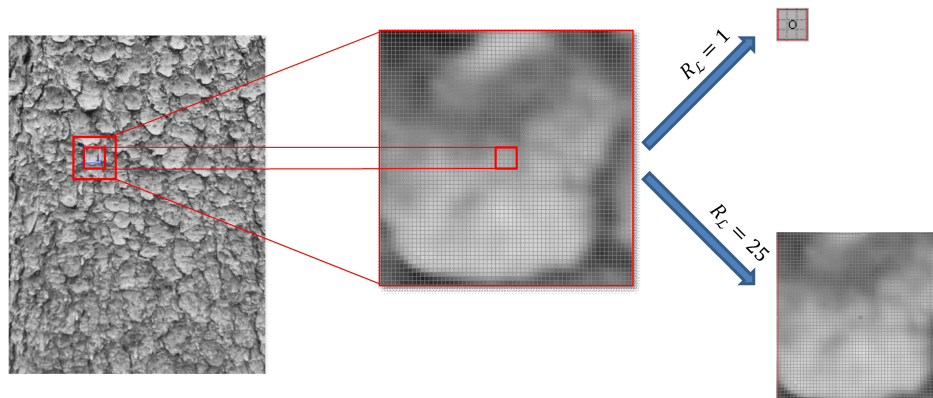


FIGURE 3.1 – Exemple de représentation d'une structure de texture d'écorce. Un large support spatial permet de représenter l'information de la macrostructure, tandis qu'un support très local ne le permet pas.

3.1 Modèle de description structurale

L'approche proposée pour la description structurale de la texture d'écorce est inspirée des motifs binaires LBP. En effet, de façon analogue au descripteur LBP qui encode les micro-motifs à partir de régions locales, on cherche à introduire un descripteur qui encode les macro-motifs à partir de grandes régions, i.e., des macrostructures. Spécifiquement, cette description passe par les étapes suivantes :

1. L'extraction de la macrostructure : on procède par une représentation multi-échelles selon une nouvelle conception de l'espace d'échelle afin de représenter l'information d'intensité répartie dans le support qui englobe la macrostructure.
2. La représentation de la macrostructure : à cette étape, l'information de texture contenue dans la macrostructure est représentée par des attributs intrinsèques afin de dériver une représentation réduite ayant certaines propriétés d'invariance.
3. Encodage de la macrostructure : les variations d'information d'intensité dans la macrostructure sont encodées selon un schéma binaire pour produire un code macro-motif caractéristique de la macrostructure.
4. Le descripteur de texture d'écorce : à cette étape, on génère le descripteur de l'image à partir de la distribution des codes macro-motifs. Ce descripteur peut être calculé à différents niveaux de représentation, i.e, globale, par région, ou bien locale.

De façon plus formelle, la description de la texture d'écorce est modélisée par le concept de composition d'applications. Une application permet de réaliser une opération de projection de l'information d'un espace de dimension élevée D vers un espace de taille réduite $|E|$ avec $|E|$ inférieure ou égale à $|D|$, permettant ainsi de générer une nouvelle représentation compacte avec des propriétés intrinsèques importantes.

Soit l'image en entrée \mathcal{I} , un pixel donné x_c de l'image \mathcal{I} est représenté par un code macro-motif avec \mathcal{L} niveaux d'échelle selon un processus de calcul basé sur la composition de trois applications, comme suit :

$$\begin{aligned}
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E} : \mathbb{Z} &\rightarrow \mathbb{Z} \\
 x_c &\mapsto \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \mathcal{C}(\mathcal{R}(\mathcal{E}(x_c))) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \text{macro-motif}
 \end{aligned} \tag{3.1}$$

où

- \mathcal{E} est l'application qui permet l'extraction de la macrostructure autour de x_c .
- \mathcal{R} est l'application qui permet la représentation de l'information contenue dans la macrostructure.
- \mathcal{C} est l'application qui permet l'encodage de l'information représentée dans la macrostructure.

Il est à noter que cette modélisation, par le principe de combinaison d'applications, généralise la représentation des microstructures par les micro-motifs LBP à partir d'une image monochrome avec un seul niveau d'échelle $\mathcal{L} = 1$. En effet, on variant les espaces de départ et arrivée de chacune

de ces applications on retrouve le cas spécifique du LBP avec un voisinage local $\mathcal{X}_{1,8}$ d'un rayon $R = 1$ et $N = 8$ points voisins comme suit :

$$\begin{aligned}
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E} : \mathbb{Z} &\rightarrow [0, 2^8 - 1] \\
 x_c &\mapsto \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \mathcal{C}(\mathcal{R}(\mathcal{E}(x_c))) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \mathcal{C}(\mathcal{R}(\mathcal{X}_{1,8})) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \mathcal{C}(\{x_n\}) \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \sum_{n=1}^8 (x_n - x_c) 2^{n-1} \\
 \mathcal{C} \circ \mathcal{R} \circ \mathcal{E}(x_c) &= \text{micro-motif}
 \end{aligned} \tag{3.2}$$

où l'application \mathcal{E} permet d'extraire le voisinage $\mathcal{X}_{1,8}$. Ce voisinage est représenté par les valeurs d'intensité en niveaux de gris des x_n via l'application \mathcal{R} . Finalement, l'application \mathcal{C} encode l'information dans ce voisinage par un gradient directionnel qui reflète la relation entre les voisins x_n et le pixel central x_c par un code *micro-motif*. La représentation résultante a une propriété d'invariance aux changements monotoniques du niveau de gris [OJALA et collab. \[2002\]](#).

3.2 Extraction de la macrostructure

La première étape pour la description de la texture d'écorce est l'extraction de la macrostructure autour d'un pixel donné x_c dans l'image \mathcal{I} . L'idée ici est de définir une conception de l'espace d'échelle pour extraire un ensemble de résolutions spatiales, i.e., les niveaux d'échelle de la macrostructure. Ainsi, une macrostructure est un ensemble de \mathcal{L} niveaux d'échelle $\{\mathcal{X}_{R_i, N_i} \mid i \in \llbracket 1, \mathcal{L} \rrbracket\}$ (c.f. Figure 3.2).

L'application \mathcal{E} est réécrite comme suit :

$$\begin{aligned}
 \mathcal{E} : \mathbb{Z} &\rightarrow \mathbb{R}^{\mathcal{L}} \\
 x_c &\mapsto \mathcal{E}(x_c) \\
 \mathcal{E}(x_c) &= \{\mathcal{X}_{R_i, N_i} \mid i \in \llbracket 1, \mathcal{L} \rrbracket\}
 \end{aligned} \tag{3.3}$$

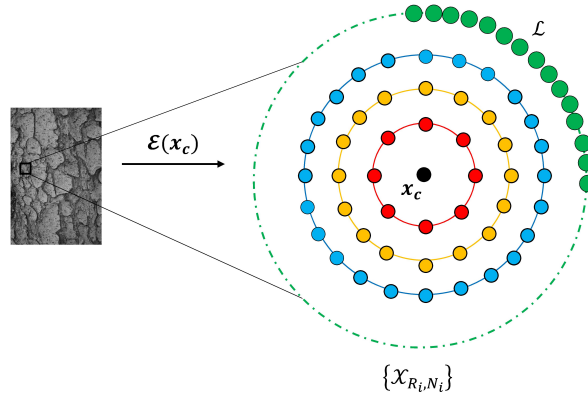
où : \mathcal{X}_{R_i, N_i} est le $i^{\text{ème}}$ niveau d'échelle caractérisé par un vecteur d'intensité de N_i points voisins, i.e., $N_i = [x_n]^T$, $|n \in \llbracket 1, N_i \rrbracket$ situé à une distance R_i du pixel central x_c .

Par conséquent, une macrostructure est caractérisée par :

- Une distribution des rayons R_i .
- Une distribution des points voisins N_i .

3.2.1 Distribution des niveaux d'échelle

Le paramètre \mathcal{L} régularise la taille de la macrostructure en terme de nombre d'échelles. Plus la valeur de \mathcal{L} est grande, plus la macrostructure s'étale sur un support spatial plus large, ce qui permet de garder une information distante de la texture.


 FIGURE 3.2 – Combinaison d’applications pour l’extraction d’une macrostructure de \mathcal{L} niveaux.

Selon l’illustration présentée dans la Figure 3.1, la macrostructure de l’écorce est définie à des niveaux d’échelle relativement élevés, et peut être modélisée par une combinaison de \mathcal{L} niveaux concentriques autour de x_c .

La nouvelle conception de l’espace d’échelle repose sur la définition d’une *distance radiale*, en pixels, entre les niveaux d’échelle. Cette distribution est en partie inspirée par la représentation de l’espace d’échelle des variantes LBP multi-échelles : OJALA et collab. [2002] propose une distribution linéaire dense limitée à seulement trois niveaux d’échelle. Ce modèle a été appliqué dans diverses variantes LBP multi-échelles LIU et collab. [2017]. La distribution proposée dans ce travail est différente de l’existant. En effet, on cherche à étendre le support spatial de la macrostructure afin d’introduire d’avantage de détails de texture à partir des pixels distants.

Soit une macrostructure de taille \mathcal{L} , on exprime la distribution de ces niveaux d’échelle comme suit :

$$R_i = R_{i-1} + \tau; \quad R_1 = 1, \quad i \in \llbracket 1, \mathcal{L} \rrbracket. \quad (3.4)$$

Ceci implique que la distance entre une paire de niveaux d’échelle adjacents ($\mathcal{X}_{R_{i+1}, N_{i+1}}, \mathcal{X}_{R_i, N_i}$) est toujours égale à une valeur de déplacement radial τ .

Le paramètre τ régularise la densité de la résolution spatiale. En effet, une faible valeur de τ correspond à une distribution fine (*fine*) avec des niveaux d’échelle consécutifs assez proches, donnant une forte densité de la résolution spatiale. Cela permet d’encoder les détails de l’information de la texture. En revanche, une valeur élevée de τ correspond à une faible densité de la résolution spatiale, donc une représentation grossière (*coarse*) de l’espace d’échelle. Ici, les niveaux d’échelle consécutifs sont assez distants les uns des autres, et cela permet d’encoder une information de texture répartie sur un large support spatial (c.f Figure 3.3).

3.2.2 Distribution des points voisins

En plus de son emplacement dans l’espace d’échelle, chaque niveau d’échelle \mathcal{X}_{R_i, N_i} porte une information sur le nombre de points voisins N_i et leur distribution, ainsi que leurs valeurs d’intensité afin de déterminer la quantité d’intensité à utiliser pour caractériser la macrostructure. Les points voisins N_i constituent le voisinage multi-échelles réparti autour du pixel central x_c . C’est à partir des valeurs d’intensité de ces points voisins qu’on arrive à tirer une représentation

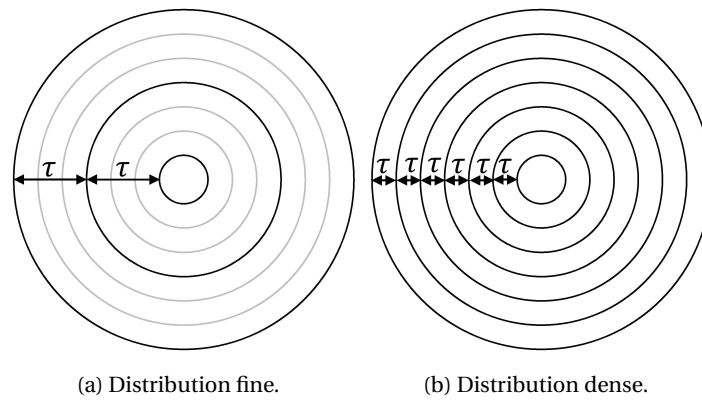


FIGURE 3.3 – Distribution des niveaux de l'espace d'échelle.

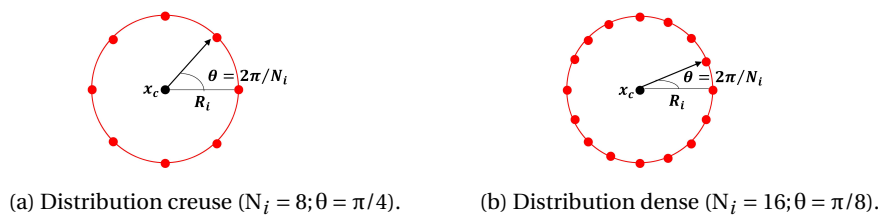


FIGURE 3.4 – Distribution des points voisins.

de la macrostructure.

La distribution des N_i points voisins est régularisée par un paramètre θ qui exprime la distance circonférentielle entre deux points successifs dans le $i^{\text{ème}}$ niveau d'échelle, tel que $\theta = 2\pi/N_i$. Plus précisément, en fonction du nombre de points voisins N_i , on peut modéliser une représentation du voisinage allant de creuse à dense (*sparse to dense*). En effet, une faible valeur de points voisins N_i correspond à une grande valeur de déplacement circonférentiel θ , et reflète une représentation creuse de l'information autour du pixel central x_c (c.f Figure 3.4a). Par contre, un nombre élevé de voisins N_i correspond à une faible valeur de θ et produit une représentation dense de l'information autour du point central x_c (c.f Figure 3.4b).

De façon générale, la densité du voisinage est exprimée en fonction de l'angle θ qui représente aussi une information directionnelle relative à la structure de la texture, comme le montre la Figure 3.4. Le cas spécifique du voisinage LBP à huit points voisins présente une distribution selon les quatre directions principales, à savoir $\theta = 2\pi/8 : \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ OJALA et collab. [2002].

Concernant les valeurs d'intensité des points voisins N_i au $i^{\text{ème}}$ niveau d'échelle, on retrouve que lors de l'acquisition, l'image brute de l'écorce est représentée dans l'espace de couleurs primaires RGB. Cependant, l'espace RGB à la particularité d'être sensible aux changements de luminosité. Par conséquent, et afin de réduire cette sensibilité, il convient de produire une nouvelle représentation des valeurs d'intensité. Ce nouvel espace peut être le niveau de gris ou bien un plan d'un espace de couleur donné où l'information de luminosité est séparée de l'information de chrominance, comme par exemple le plan Teinte (Hue H) de l'espace HSV, le plan Luminance (Luma Y) de l'espace YCbCr, ou bien encore le plan L de l'espace $L^* a^* b$. Ainsi, selon l'espace d'intensité choisi, une nouvelle valeur d'intensité est calculée pour les points voisins N_i ainsi que le pixel central x_c .

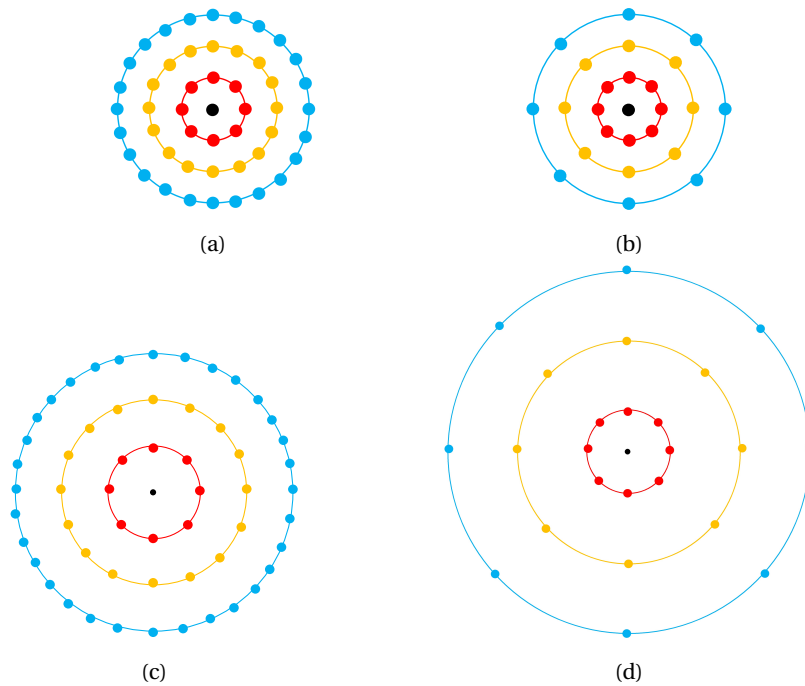


FIGURE 3.5 – Exemple de différentes macrostructures avec $\mathcal{L} = 3$ niveaux. (a) Distribution dense pour les niveaux d'échelle et des points voisins. (b) Distribution dense pour les niveaux d'échelle et creuse pour des points voisins. (c) Distribution grossière pour les niveaux d'échelle et dense des points voisins. (d) Distribution grossière pour les niveaux d'échelle et creuse des points voisins.

Au final, en fonction du nombre d'échelles et leur distribution, et du nombre de points voisins et leur distribution, différents schémas de l'espace d'échelle peuvent être générés. La Figure 3.5 montre quelques représentations d'une macrostructure de $\mathcal{L} = 3$ niveaux avec une représentation dense pour les niveaux d'échelle et des points voisins (c.f Figure 3.5a), une représentation dense au niveau spatial, mais creuse pour les points voisins (c.f Figure 3.5b), une représentation grossière au niveau spatial mais dense au niveau voisinage (c.f Figure 3.5c), et enfin, une représentation grossière au niveau spatial et creuse au niveau voisinage (c.f Figure 3.5d).

3.3 Représentation de la macrostructure

Suite à l'extraction de la macrostructure, un voisinage multi-échelles est défini autour d'un pixel donné x_c dans l'image \mathcal{I} , et englobe une information structurale de la texture. Cependant, encoder directement cette information n'est pas avantageux en raison de ce qui suit :

- ▶ Le nombre important de points voisins à des niveaux d'échelle élevés ralentit le traitement. En effet, une représentation efficace de la macrostructure requiert un nombre important de points voisins. Ce nombre s'accroît en fonction de la taille de l'espace d'échelle. Une représentation réduite de ce voisinage s'avère nécessaire.
- ▶ Plus le voisinage est grand, plus l'effet de repliements (*aliasing*) est récurrent, i.e., les valeurs d'intensité des points voisins distants sont souvent bruitées MÄENPÄÄ et PIETIKÄINEN [2003]. Par conséquent, de nombreuses variantes appliquent un mécanisme de tolérance au bruit via le calcul de patch (i.e., une région centrée autour de chacun des points voisins

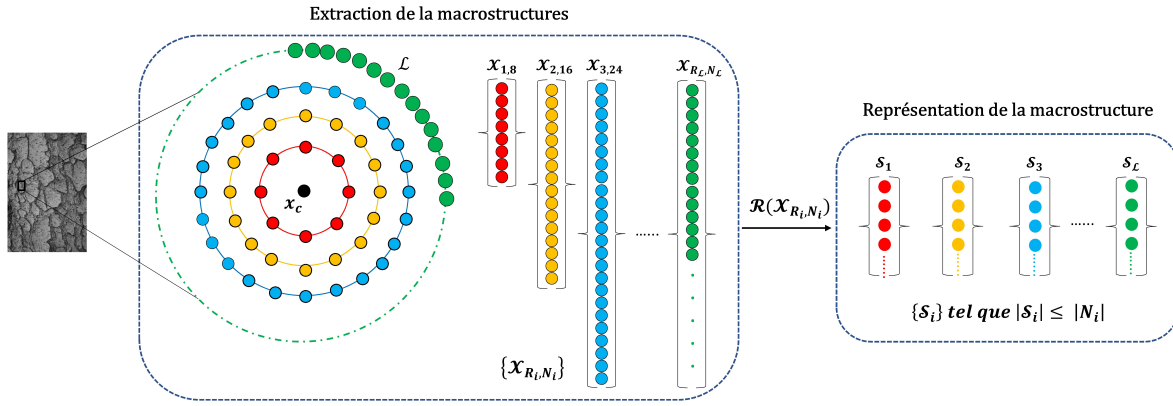


FIGURE 3.6 – Représentation des niveaux d'échelle de la macrostructure.

distants qui sert de support à un filtre passe-bas pour l'estimation de la valeur d'intensité), comme le patch moyen LIAO et collab. [2007]; WANG et collab. [2013]; WOLF et collab. [2008], le patch Gaussien MÄENPÄÄ et PIETIKÄINEN [2003]; ŠULC et MATAS [2013], ou le patch médian LIU et collab. [2016b]. Cependant, malgré l'efficacité de cette approche, elle reste gourmande en calcul, surtout lors de la convolution avec des filtres de tailles différentes d'un niveau de résolution à un autre. De ce fait, on cherche à réduire l'effet de bruit au niveau de la macrostructure sans passer par l'application locale des filtres passe-bas pour pondérer les valeurs d'intensité des points voisins distants.

- Encoder directement les valeurs d'intensité des points voisins de la macrostructure peut être insuffisant dans le cas d'une variation de luminance, de rotation, ou d'échelle. Une information supplémentaire doit être ajoutée afin d'assurer une certaine invariance à ces problèmes photométriques. Cependant, il sera plus intéressant de pouvoir extraire ses propriétés directement à partir des valeurs d'intensité des points voisins, i.e., des propriétés intrinsèques adaptatives à chaque macrostructure.

L'idée sous-jacente à notre approche est de générer une représentation réduite et homogène des niveaux de l'espace d'échelle pouvant représenter des caractéristiques intrinsèques de l'information d'intensité dans ce voisinage. Différentes représentations peuvent être extraites via l'application \mathcal{R} qui projette chaque niveau de l'espace d'échelle individuellement dans un nouvel espace, et retourne une représentation réduite. L'application \mathcal{R} est donnée par la formule suivante :

$$\begin{aligned} \mathcal{R} : \mathbb{R}^{N_i} &\rightarrow \mathbb{R}^M \\ \mathcal{X}_{R_i, N_i} &\mapsto \mathcal{R}(\mathcal{X}_{R_i, N_i}) \\ \mathcal{R}(\mathcal{X}_{R_i, N_i}) &= \mathcal{S}_i \end{aligned} \quad (3.5)$$

où \mathcal{S}_i est la $i^{\text{ème}}$ représentation du niveau d'échelle \mathcal{X}_{R_i, N_i} dans une macrostructure centrée autour de x_c .

Plus explicitement, pour chaque niveau \mathcal{X}_{R_i, N_i} , on projette le vecteur d'intensité $[x_m]^T$ de dimension N_i dans un espace de dimension $M | M \leq N_i$ pour obtenir un nouveau vecteur représentatif \mathcal{S}_i , tel que : $\mathcal{S}_i = [s_m]^T \quad | m \in \llbracket 1, M \rrbracket; \quad i \in \llbracket 1, \mathcal{L} \rrbracket$.

Suite à l'application de l'opération de projection dans l'espace de représentation via l'application \mathcal{R} pour chacun des \mathcal{L} niveaux d'échelle de la macrostructure, on obtient une représentation

réduite et uniforme pour tous les niveaux d'échelle (c.f Figure 3.6).

Une autre façon de procéder consiste à projeter l'ensemble des niveaux de la macrostructure (c'est à dire, considérer tous les vecteurs N_i à la fois et non pas individuellement), donnés par un vecteur de dimension $N_i \times \mathcal{L}$, dans un nouvel espace, et générer un seul vecteur représentatif S de dimension P . Toutefois, malgré que cette démarche produit une représentation remarquablement compacte de l'information contenue dans la macrostructure, elle ne préserve pas l'information spatiale propre à chaque niveau de l'espace d'échelle. En effet, à chaque niveau de la macrostructure, une relation de voisinage existe entre les points voisins de ce niveau relatif à la résolution de la structure de texture, or suivant cette façon de faire, l'information par résolution spatiale sera perdue. À cet effet, l'approche de représentation introduite dans cette partie est considérée comme préservatrice de l'information du voisinage par résolution spatiale.

3.4 Encodage de la macrostructure

À ce stade, pour une image \mathcal{I} , autour d'un pixel donné x_c on a pu extraire une macrostructure (ou bien un voisinage multi-échelles de \mathcal{L} niveaux), et compacter l'intensité à chaque niveau de l'échelle via une projection dans un espace de taille réduite. Ainsi la macrostructure est représentée par \mathcal{L} vecteurs S_i .

L'objectif de notre approche pour la description de la texture d'écorce est de produire un seul code macro-motif pour la macrostructure. Retenir la représentation $\{S_i\}$, $|i \in [1, \mathcal{L}]$ comme code macro-motif n'est pas une solution optimale. En effet, cela est proportionnel au nombre de macrostructures extraites de l'image \mathcal{I} . Dans le cas où une macrostructure de taille \mathcal{L} est extraite pour chaque pixel dans l'image, le descripteur final de l'image sera d'une dimension considérable, i.e., $Dim = S \times \mathcal{L} \times H \times W$ (où H et W sont la hauteur et la largeur de l'image \mathcal{I}).

De ce fait, une étape d'encodage est nécessaire pour compacter d'avantage l'information dans la macrostructure. Pratiquement, à cette étape, on cherche à encoder la relation entre les niveaux d'échelle de la macrostructure représentés par les vecteurs S_i . Ceci revient à encoder les variations de l'information structurale pour produire un code caractéristique de la macrostructure.

Le schéma classique d'un encodage multi-échelles, MResLBP [OJALA et collab. \[2002\]](#), caractérise la relation entre chaque niveau de la résolution et le pixel central x_c par un code LBP. Cependant, ce schéma a les inconvénients suivants :

- ▶ L'information inter-échelles (ou bien la relation entre les niveaux d'échelle) dans la macrostructure est complètement négligée, du fait que chaque niveau d'échelle est encodé indépendamment des autres niveaux dans l'espace d'échelle. Ceci conduit à une perte importante de l'information structurale.
- ▶ Le descripteur final de la texture est une combinaison de plusieurs descripteurs LBP par résolution, que ce soit par une concaténation ou par une représentation conjointe ne produit pas un code unique pour représenter la macrostructure. Par conséquent, la dimension finale du descripteur de l'image de texture est relativement élevée.

D'autres variantes LBP ont introduit des solutions afin de remédier aux limites de ce schéma en encodant une information inter-échelles [LIU et collab. \[2012a\]](#); [QI et collab. \[2013, 2014\]](#). Ce-

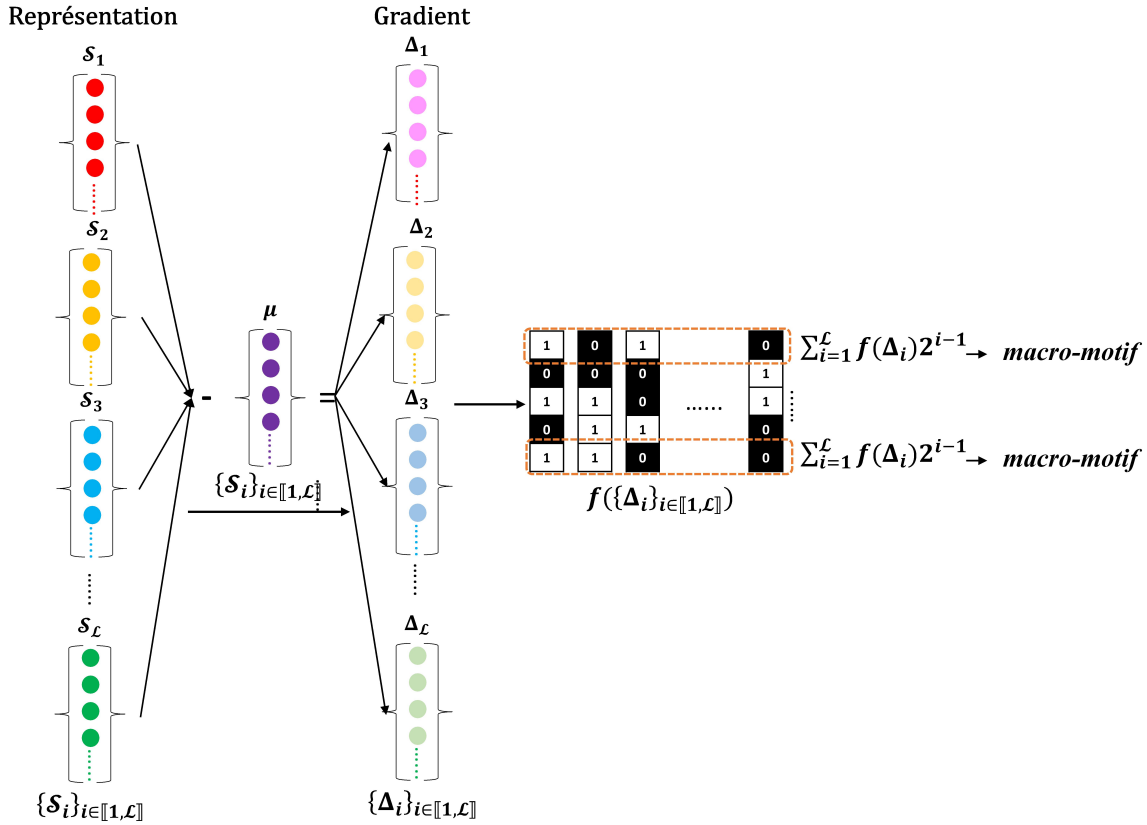


FIGURE 3.7 – Combinaison d’applications pour l’encodage de la macrostructure et construction des *macro-motifs*.

pendant, cela reste insuffisant, car pour ces variantes, le code qui représente la macrostructure reste une combinaison de codes calculés par niveau d’échelle.

Différemment de ces approches, notre schéma d’encodage vise à produire un seul code *macro-motif* pour représenter la macrostructure, tout en exploitant les variations d’information entre les niveaux d’échelle.

Dans le contexte de la modélisation par combinaison d’application, on formalise la relation entre le niveau d’échelle de la macrostructure par l’application \mathcal{C} (c.f Figure 3.7). Cette dernière est définie comme suit :

$$\begin{aligned}
 \mathcal{C} : \mathbb{R}^{\mathcal{L}} &\rightarrow \llbracket 0, 2^{\mathcal{L}} - 1 \rrbracket \\
 \{\mathcal{S}_i\} &\mapsto \mathcal{C}(\{\mathcal{S}_i\}) \\
 \mathcal{C}(\{\mathcal{S}_i\}) &= \text{macro-motif}, \quad |i \in \llbracket 1, \mathcal{L} \rrbracket
 \end{aligned}
 \tag{3.6}$$

D’un point de vu traitement de texture, l’encodage d’une variation d’information permet de représenter des structures avec un changement d’intensité notable comme les coins, les contours, etc. Ceci revient à effectuer un calcul de gradient, ou bien une différence signée qui exprime le degré de variation de l’information. Cette différence est composée de magnitude et de signe. Cependant, le signe de la différence a été utilisé dans de nombreuses variantes LBP pour caractériser la variation de l’information [OJALA et collab. \[2002\]](#); [TAN et TRIGGS \[2010\]](#). Bien que la magnitude a permis d’ajouter une information complémentaire et d’améliorer la performance de certains descripteurs comme dans [GUO et collab. \[2010a\]](#); [LIU et collab. \[2016b, 2014\]](#); [ZHAO et collab. \[2012b\]](#), son utilisation dépend du domaine d’application et des bases d’images utilisées [LIU et collab.](#)

[2016a]. De ce fait, on caractérise la variation de l'information tirée par le calcul de gradient en utilisant seulement le signe de la différence.

Le calcul du gradient consiste à encoder la relation entre les niveaux d'échelle de la macro-structure, représentés par les vecteurs \mathcal{S}_i via le calcul d'une différence entre chacun de ces vecteurs et un prototype adaptatif μ au lieu du pixel central x_c (comme le cas de l'encodage LBP). En effet, l'utilisation du pixel central comme seuil de binarisation produit un code sensible aux moindres fluctuations dans la valeur de x_c ce qui peut générer des représentations différentes du *macro-motif*. Par conséquent, un seuil adaptatif calculé à partir des représentations des niveaux de la macrostructure est plus stable contre les éventuels changements des valeurs d'intensité.

Le calcul du gradient Δ_i est donné par la formule suivante :

$$\Delta_i = \mathcal{S}_i - \mu, \quad i \in \llbracket 1, \mathcal{L} \rrbracket \quad (3.7)$$

Ainsi, on génère Δ_i différences signées relatives à chaque niveau de l'espace d'échelle, et dont le signe caractérise les variations de l'intensité dans la macrostructure d'un niveau d'échelle à un autre.

Pour le calcul du code *macro-motif*, les valeurs signées du gradient Δ_i sont quantifiées via une fonction f . Les niveaux de quantification ainsi que les seuils restent des paramètres à déterminer. Cependant, d'un côté, des niveaux de quantification élevés caractérisent différents aspects de la variation de l'information et réduisent la sensibilité au bruit NANNI et collab. [2010]; TAN et TRIGGS [2010]. De l'autre côté, le choix et le calcul des seuils de quantification sont relatifs au domaine d'application ainsi qu'aux caractéristiques des bases d'images testées. De ce fait, on procède par une quantification binaire, i.e, une binarisation des valeurs de gradient par rapport à un seul seuil μ dont la valeur est déterminée de manière adaptative.

Le calcul final pour générer le code *macro-motif* est réalisé par la formule suivante :

$$\begin{aligned} \text{macro-motif} &= \sum_{i=1}^{\mathcal{L}} f(\Delta_i) 2^{i-1} \\ f(\Delta_i) = f(\mathcal{S}_i - \mu); \quad f(\mathcal{S}_i - \mu) &= \begin{cases} 1 & \text{si } \mathcal{S}_i \geq \mu \\ 0 & \text{sinon.} \end{cases} \end{aligned} \quad (3.8)$$

Ainsi, via ce processus, un pixel donné x_c dans l'image \mathcal{I} est représenté par un code *macro-motif* calculé à partir d'un voisinage multi-échelles centré autour de lui. En plus des propriétés d'invariance tirées des représentations \mathcal{S}_i , le code macro-motif présente une stabilité face aux valeurs bruitées, et une invariance aux changements monotoniques du niveau de gris, car basé essentiellement sur un encodage LBP.

3.5 Descripteur de la texture d'écorce

La construction du descripteur final d'une image de texture d'écorce est essentiellement basée sur l'encodage d'attributs informatifs et pertinents. En plus de la capacité de modéliser la diversité d'apparence des espèces végétales, l'efficacité du descripteur de la texture d'écorce est aussi mesurée par son pouvoir discriminant en vue d'une éventuelle similarité inter-espèces.

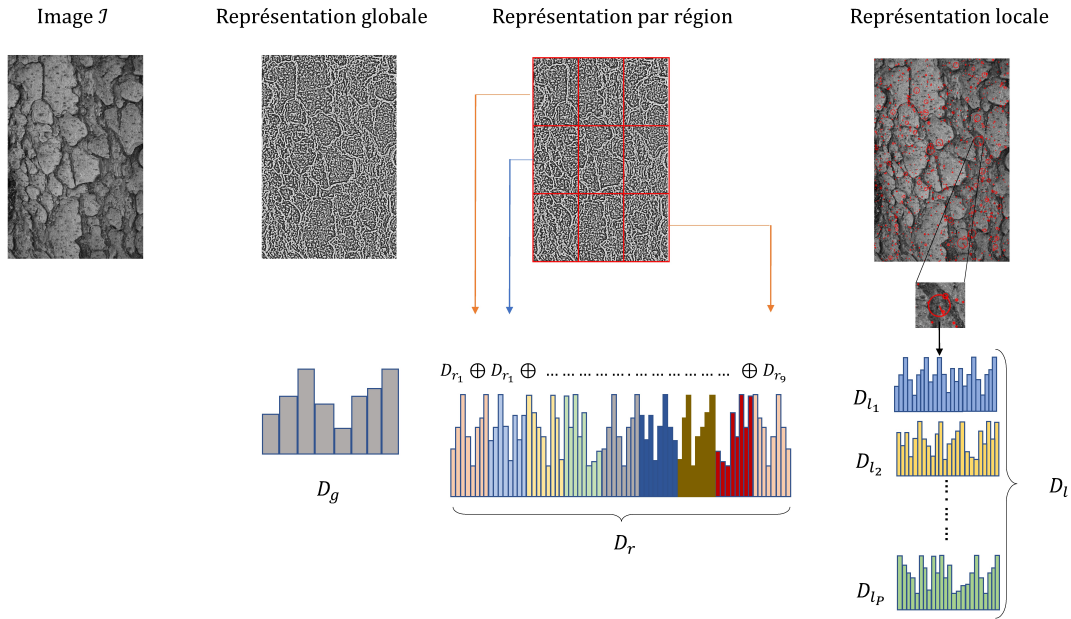


FIGURE 3.8 – Représentation de la texture selon différents niveaux de descriptions : global, par région, et local par des points d'intérêt.

À ce niveau, les pixels de l'image \mathcal{I} sont caractérisés par des codes macro-motifs. La distribution de ces macro-motifs est représenté par un histogramme de $2^{\mathcal{L}}$ bins, où chaque macro-motif vote dans le bin correspondant à sa valeur décimale. Ainsi, le descripteur final de l'image est une représentation de la distribution des codes macro-motifs. Ceci résume d'autant plus l'information tout en produisant une caractérisation compacte, à la fois structurale (les codes macro-motifs), et statistique (la distribution de ces codes).

Toutefois, la construction du descripteur de l'image \mathcal{I} peut se faire à différents niveaux de représentation, à savoir, globale, par région, ou bien locale (c.f Figure 3.8). On désigne par le niveau de représentation, l'approche empruntée pour le calcul des codes macro-motifs dans le sens où le processus est appliqué sur tout l'ensemble des pixels de l'image \mathcal{I} ou bien sur certaines régions ayant des propriétés spatiales (l'emplacement spatial de régions plus importantes selon le contenu de l'image), ou des propriétés structurales définies par des changements d'intensité (des bords, des coins, etc).

3.5.1 Représentation globale

Une description de l'image est dite globale, D_g , si le même principe d'encodage est appliqué pour tous les pixels de l'image sans tenir compte de certaines caractéristiques de leurs emplacements spatiaux ou leur valeur d'intensité.

Pour une image \mathcal{I} de taille $H \times W$, en appliquant un encodage donné, chaque pixel de cette image est représenté par un code. Ceci est équivalent à la génération d'une image de codes \mathcal{I}_{code} à partir de l'image \mathcal{I} ayant la même taille $H \times W$. Le descripteur final est un histogramme qui représente la distribution de ces codes. Le calcul de cet histogramme est donné par :

$$D_g = \sum_{h=1}^H \sum_{w=1}^W \mathcal{I}_{code}(h, w) = c \quad | \quad c = 0, 1, 2, \dots, C-1 \quad (3.9)$$

où C est le nombre total des codes binaires possibles.

Suivant notre encodage, \mathcal{I}_{code} est l'image des macro-motifs et les valeurs possibles de macro-motifs sont au nombre de $C = 2^{\mathcal{L}}$.

Cette description, dite dense, a l'avantage d'encoder toute l'information de texture de l'image de l'entrée et produire une représentation fiable de toutes les structures de la texture.

3.5.2 Représentation par région

À cette échelle de représentation, l'image en entrée est divisée en K régions selon une grille superposée sur la surface de l'image. Ensuite, le même principe d'encodage est appliqué sur toutes les régions de l'image produisant un descripteur par région. La représentation finale de l'image est ainsi la combinaison (\oplus) de tous ces descripteurs régionaux. Formellement, le calcul du descripteur par région D_r est donné par :

$$D_r = \oplus_{k=1}^K (w_k D_{r_k}); \quad \sum_{k=1}^K w_k = 1 \quad (3.10)$$

où w_k est le poids de la région k , et D_{r_k} est le descripteur de la région k . Il est calculé par la formule 3.9. Dans ce cas, une description par région génère un descripteur de dimension : $K \times 2^{\mathcal{L}}$.

Une description par région a la propriété d'être robuste à la translation. Cependant, cette description nécessite une connaissance à priori du partage des régions sur la grille. On retrouve cette description le plus souvent dans des applications de reconnaissance faciale : l'image de visage est d'abord alignée puis divisée en régions (yeux, nez, bouche, front, etc) bien déterminées et qu'on retrouve dans toutes les images faciales. Cependant, dans le contexte des images d'écorce, l'information de la structure de la surface n'est pas répertoriée selon des régions bien délimitées. De plus, il n'y pas besoin d'une invariance à la translation dans une application d'identification des espèces végétales.

3.5.3 Représentation locale

À l'encontre de la description globale, une description locale D_l est réalisée par l'application du même principe d'encodage sur certains pixels de l'image. Ces pixels ont une propriété structurale. Ceci se caractérise généralement par un changement d'intensité désignant la présence d'une structure locale comme les coins, bords, etc.

Une description locale est le plus souvent réalisée dans des applications sur des images ou scènes naturelles qui contiennent plusieurs objets en plus de l'objet d'intérêt. Procéder ainsi permet la construction d'un descripteur de l'image sans passer par une étape de segmentation, pour isoler l'objet d'intérêt de son fond. Ainsi dit, il est nécessaire de déterminer l'ensemble des pixels qui seront encodés. Ces pixels sont communément appelés des points d'intérêt (*keypoints*), et sont détectés par des détecteurs de points d'intérêt comme Harris HARRIS et STEPHENS [1988], SIFT LOWE [2004], SURF BAY et collab. [2006], etc.

Selon cette représentation, la description finale de l'image \mathcal{I} est l'ensemble de tous les descripteurs des régions autour des P points d'intérêt détectés dans l'image \mathcal{I} :

$$D_l = \{D_{l_p}\}; \quad (3.11)$$

où D_{I_p} est le descripteur du $p^{\text{ème}}$ point d'intérêt. Dans le cas de notre approche, une description locale avec par exemple des points d'intérêt SIFT, est de dimension : $P \times 2^{\mathcal{L}}$.

Lors de l'étape de classification, on peut procéder soit à une mise en correspondance (*matching*) des points d'intérêt d'une paire d'images, soit à générer une description globale par le modèle (*Bag of Words*(BoW)) CSURKA et collab. [2004] afin de simplifier la représentation du descripteur final.

3.6 Conclusion

Dans ce chapitre, nous avons présenté la modélisation de l'approche proposée pour la représentation de la texture d'écorce. Cette approche est formalisée à l'aide de la composition d'applications qui permettent de réaliser une opération de projection dans un nouvel espace de taille réduite et finie, pouvant apporter des propriétés d'invariance. L'intérêt de notre approche est de construire un code macro-motif pour la description d'un voisinage multi-échelles autour d'un pixel donné dans l'image d'écorce.

Cette formalisation générique de l'approche proposée permet de dériver une famille de descripteurs macro-motifs. Ceci est réalisable via la variation des espaces de projection à chaque étape du traitement, ce qui permet de représenter différents aspects de la structure d'écorce, et contribuer à une caractérisation efficace de la macrostructure spécifique des différentes espèces d'écorce. Les variantes proposées seront exposées avec plus de détails dans le chapitre suivant.

Chapitre 4

Descripteurs des macro-motifs binaires statistiques

Sommaire

4.1 Conception de la macrostructure	64
4.1.1 Espace d'échelle	64
4.1.2 Voisinage multi-échelles	67
4.2 Distribution d'intensité	69
4.2.1 Mesures Statistiques	70
4.2.2 Représentation statistique	74
4.3 Codes macro-motifs	76
4.3.1 Macro-motif binaire statistique radial	77
4.3.2 Macro-motif binaire statistique profond	77
4.4 Descripteurs de la texture d'écorce	81
4.4.1 SRBP : <i>Statistical Radial Binary Pattern</i>	81
4.4.2 SMBP : <i>Statistical Macro Binary Pattern</i>	82
4.5 Propriétés des codes macro-motifs	82
4.6 Conclusion	84

UNE fois la modélisation de notre approche pour la description de la texture d'écorce établie, on est en mesure d'introduire des spécificités sur chacune des étapes du calcul pour obtenir une famille de descripteurs de la texture d'écorce, qui, communément décrivent le motif caractéristique des différentes structures d'écorce mais encodent différents aspects de cette structure.

La description du contenu visuel des surfaces d'écorce doit permettre une représentation fiable des éléments caractéristiques et discriminants pouvant ainsi faciliter leur différenciation. Le processus décrit ici est essentiellement basé sur l'encodage d'éléments intrinsèques simples, porteurs d'information structurale, compacte, et moins sensibles aux variations photométriques les plus souvent rencontrées dans des applications sur des objets naturels.

Le calcul des descripteurs de texture décrit dans cette partie est basé sur les applications présentées en Chapitre 3. Il s'agit de varier les espaces de projection des applications utilisées à chaque stade du calcul du code macro-motif, à savoir, à l'extraction, la représentation, et l'encodage de la macrostructure. Plus spécifiquement, on introduit différentes conceptions de l'espace d'échelle et du voisinage multi-échelles, différentes représentations de l'information dans la macrostructure, et deux différents encodages de gradients qui captent chacun une information propre à la distribution de l'information d'intensité dans la macrostructure, et qui contribuent à la description des divers types de structures d'écorce. Par la suite un mécanisme de combinaison de descripteurs est proposé afin d'apporter une complémentarité de l'information et enrichir la description de la texture d'écorce.

4.1 Conception de la macrostructure

L'extraction de la macrostructure autour d'un pixel donné x_c dans une image d'écorce \mathcal{I} est réalisée par l'application \mathcal{E} . Rappelons que la formule générique de \mathcal{E} est donnée par :

$$\begin{aligned} \mathcal{E} : \mathcal{Z} &\rightarrow \mathbb{R}^{\mathcal{L}} \\ x_c &\mapsto \mathcal{E}(x_c) \\ \mathcal{E}(x_c) &= \{\mathcal{X}_{R_i, N_i}\} \quad |i \in \llbracket 1, \mathcal{L} \rrbracket \end{aligned} \tag{4.1}$$

tel que l'ensemble $\{\mathcal{X}_{R_i, N_i}\}$ constitue la macrostructure autour d'un pixel x_c .

Différents choix des distributions des niveaux d'échelle et du voisinage multi-échelles peuvent ainsi être représentés par cette formule selon le besoin de l'étude.

4.1.1 Espace d'échelle

À la conception de l'espace d'échelle, la répartition des niveaux d'échelle est exprimée selon une distance τ . Cette distance est mesurée en pixels et elle exprime le pas de déplacement radial τ entre deux niveaux d'échelle consécutifs.

Les représentations de l'espace d'échelle présentées dans ce travail sont en majeure partie inspirées de la configuration isotrope de LBP [OJALA et collab. \[2002\]](#). Nous proposons d'étudier dans ce cadre deux représentations : (i) un espace d'échelle *circulaire* (*circ*), et (ii) et un espace d'échelle *en anneaux* (*ring*).

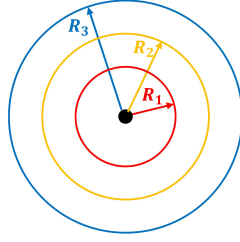
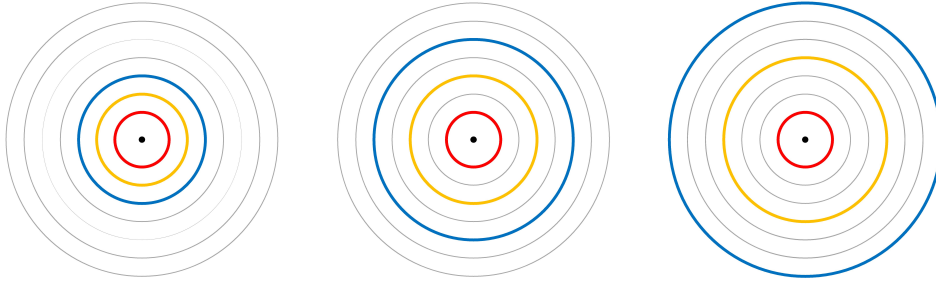
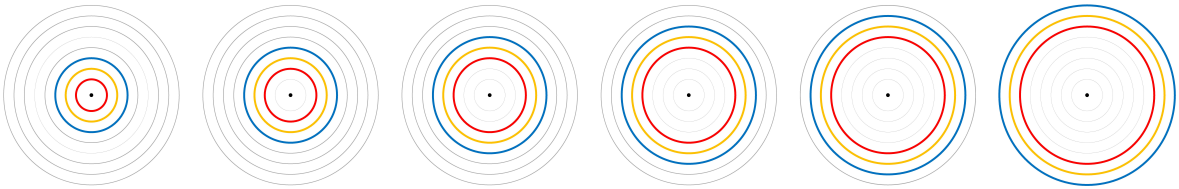


FIGURE 4.1 – Une macrostructure circulaire de trois niveaux.

FIGURE 4.2 – Distribution des niveaux d'échelle selon $\tau = \{1, 2, 3\}$.FIGURE 4.3 – Localité du voisinage avec différentes valeurs de $R_1 | R_1 \in \llbracket 1, 6 \rrbracket$, et $\tau = 1$.

Le choix d'une représentation basée sur des niveaux d'échelle circulaires, à l'encontre des niveaux d'échelle carrés, est essentiellement justifié par le fait que c'est la représentation la plus utilisée par la majeure partie des variantes LBP, notamment après son introduction par [OJALA et collab. \[2002\]](#) pour des représentations aussi bien locales que multi-échelles, à l'encontre de la version basique avec un voisinage carré.

4.1.1.1 Espace d'échelle circulaire

Étant donné une macrostructure $\{\mathcal{X}_{R_i, N_i}\}$ de \mathcal{L} niveaux. La distribution des rayons R_i des niveaux d'échelle est donnée par la formule suivante :

$$R_i = R_{i-1} + \tau; \quad i \in \llbracket 2, \mathcal{L} \rrbracket \quad (4.2)$$

Ainsi, la distance entre une paire de niveaux d'échelle ($\mathcal{X}_{R_{i-1}, N_{i-1}}, \mathcal{X}_{R_i, N_i}$) est égale à une valeur constante τ .

La Figure 4.1 montre un exemple d'espace d'échelle circulaire multi-échelles de $\mathcal{L} = 3$ niveaux. Selon la valeur de τ , on peut obtenir différentes représentations de l'espace d'échelle comme l'illustre la Figure 4.2.

Une autre façon de représenter des macrostructures encore plus larges est d'exploiter la notion de localité du voisinage. En effet, pour une macrostructure d'une certaine taille \mathcal{L} , on peut représenter une information distante en déplaçant le voisinage multi-échelles par rapport à son

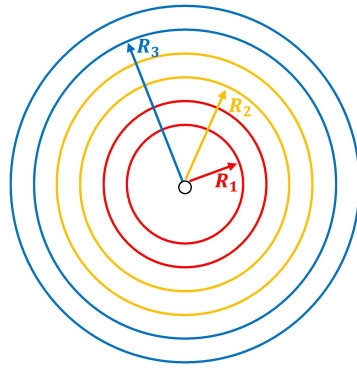


FIGURE 4.4 – Une macrostructure de trois anneaux avec deux cercles chacun.

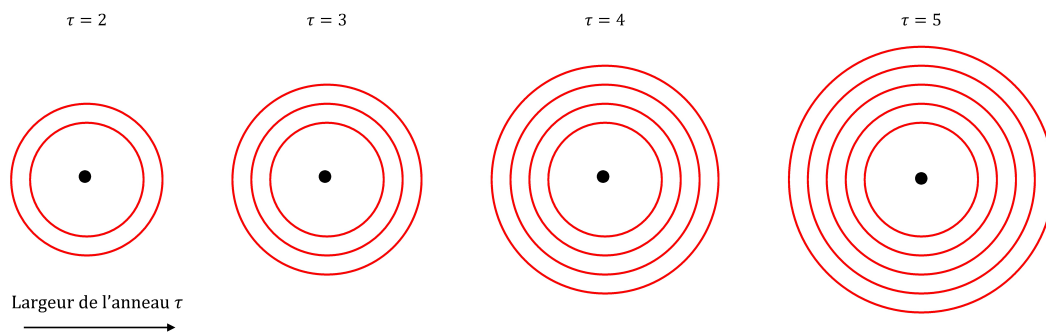


FIGURE 4.5 – Largeur d'anneau.

centre. Pratiquement, on exprime cette localité par la distance entre le premier niveau d'échelle et le pixel central x_c , qui sera simplement le premier rayon R_1

Plus les valeurs de R_1 sont grandes, plus la macrostructure est loin du pixel central x_c et donc s'étale sur un support spatial plus large. Toutefois, la taille de la macrostructure ainsi que la distribution des niveaux d'échelle reste inchangées, c'est le support spatial qui s'élargit, comme le montre la Figure 4.3.

4.1.1.2 Espace d'échelle en anneaux

Cette représentation de l'espace d'échelle est essentiellement basée sur la configuration en cercle où on considère un certain nombre de cercles consécutifs comme étant un seul niveau d'échelle (c.f. Figure 4.4). Le nombre de cercles qu'on fusionne pour définir un anneau à un niveau d'échelle de la macrostructure indique la *largeur* de l'anneau. Plus l'anneau est large, plus il contient des cercles, comme le montre la Figure 4.5 pour un niveau d'échelle avec différentes largeurs.

Formellement, une macrostructure $ring \{\mathcal{X}_{R_i, N_i}\}$ de \mathcal{L} niveaux avec \mathcal{X}_{R_i, N_i} le i^{eme} anneau avec un rayon interne R_i . Le nombre de cercles consécutifs qui forment la largeur de l'anneau est exprimé par $étendu(\llbracket R_i, R_{i+1} \rrbracket)$ tel que $R_i = R_{i-1} + \tau$, $R_1 = 1$, et τ est la distance radiale en pixel entre les rayons internes des anneaux de l'espace d'échelle.

La représentation de l'espace d'échelle en anneaux est essentiellement proposée comme solution à deux limites propres à la précédente représentation en *circ*. En effet, avec une représentation *circ* on peut voir que pour encoder une large macrostructure, on aura besoin d'un grand nombre d'échelles. Ceci a un impact direct sur la dimension de notre descripteurs qui encode es-

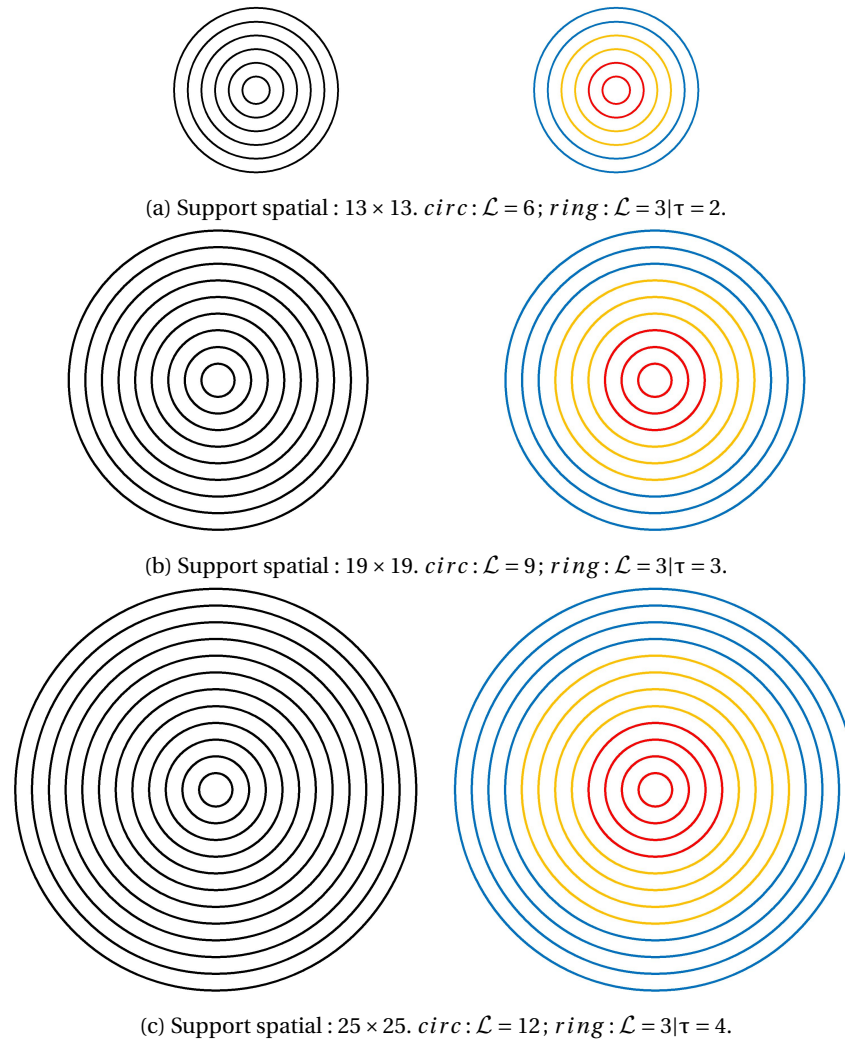


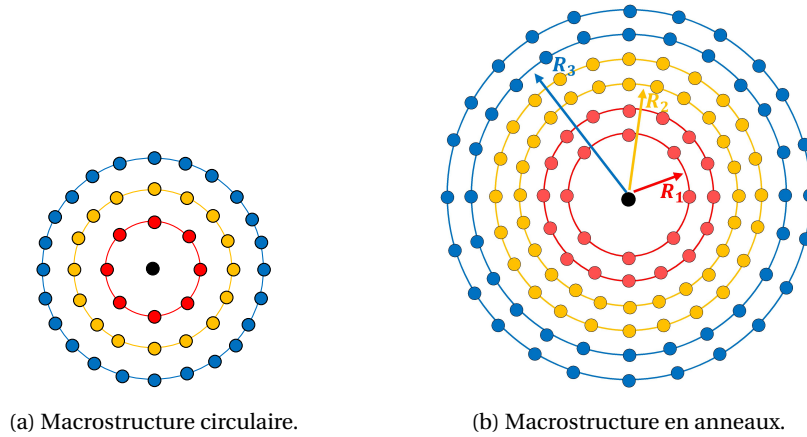
FIGURE 4.6 – Macrostructure circulaire et macrostructure en anneaux avec différentes largeurs. Exemple de comparaison en terme de taille d'espace d'échelle et support spatial pour les deux représentations de l'espace d'échelle.

sentiellement la relation inter-échelles, comme sera détaillé dans la suite de ce chapitre. De plus, une macrostructure en *circ* avec une large distance radiale ($\tau > 1$) entre les niveaux d'échelle peut générer une perte de l'information de texture. De ce fait, et afin de remédier à ces limites, on propose la représentation *ring*. Cette représentation s'étale sur un large support spatial, et ainsi on peut encoder de larges macrostructures avec un faible nombre d'échelles, et aussi préserver l'information de texture.

La Figure 4.6 illustre la correspondance entre une macrostructure *circ* et *ring* pour le même support spatial et les niveaux d'échelle \mathcal{L} correspondants. Par exemple pour un support de 19×19 une macrostructure *circ* requiert $\mathcal{L} = 12$ échelles alors qu'une macrostructure *ring* d'une largeur $\tau = 3$ représente ce support avec $\mathcal{L} = 3$ échelles seulement (c.f Figure 4.6b).

4.1.2 Voisinage multi-échelles

Un voisinage multi-échelles circulaire est donné par l'ensemble $\{\mathcal{X}_{R_i, N_i}\}, i \in [1, \mathcal{L}]$. Soit x_c le pixel central situé à l'origine $(0, 0)$, les coordonnées spatiales des N_i voisins répartis sur un cercle



(a) Macrostructure circulaire.

(b) Macrostructure en anneaux.

FIGURE 4.7 – Distribution dense des points voisins *ALLN*.

de rayon R_i , sont données par $(-R \sin(2\pi n/N), R \cos(2\pi n/N)) | n \in \llbracket 1, N_i \rrbracket$.

Au niveau de la distribution spatiale de l'espace d'échelle, on a exprimé la localité de la macrostructure ainsi que la distribution de ses niveaux d'échelle par un déplacement radial τ . De façon similaire, on définit la distribution des points voisins suivant un pas de déplacement circonférentiel θ . La distance entre les points voisins N_i au i^{eme} niveau d'échelle est exprimé par la relation $\theta = 2\pi/N_i$ OJALA et collab. [2002].

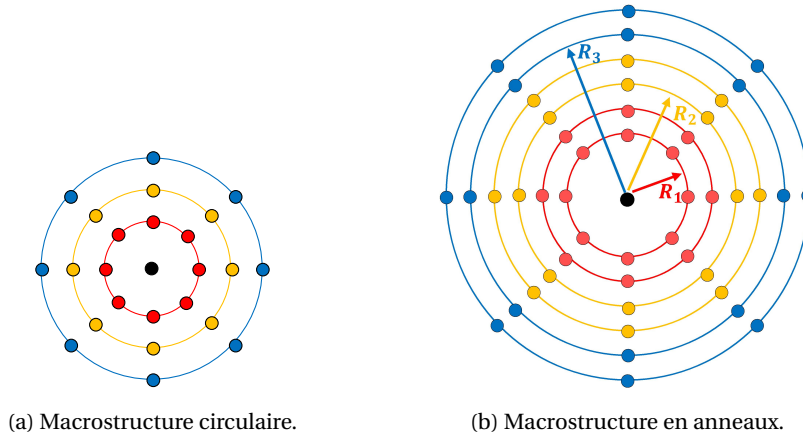
On définit deux distributions différentes des points voisins, qui expriment chacune une information sur la densité de l'information d'intensité dans le voisinage multi-échelles : une distribution dense (*ALLN*) et une distribution creuse (*OneN*).

4.1.2.1 Voisinage dense : *ALLN*

Selon une relation définie par OJALA et collab. [2002], le nombre de voisins N_i à un niveau d'échelle est défini en fonction du rayon R_i tel que $N_i = 8 \times R_i$. La valeur de l'angle θ entre les points voisins N_i devient de plus en plus petite pour des niveaux d'échelle élevés avec des cercles de rayon grand, ce qui reflète une distribution dense du voisinage.

Pour une macrostructure *ring* le nombre total des points voisins N_i est la somme de tous les points voisins des cercles qui constituent les anneaux de l'espace d'échelle. Ainsi on note par $N_{i,j}, j \in \llbracket R_{i-1}, R_i \rrbracket$ le nombre de points voisins dans le i^{eme} niveau d'échelle. Différemment d'une macrostructure *circ*, on présente une distribution *ALLN* comme suit : $N_{i,j} = 8 \times R_{i,j}$ tel que $R_{i,j} < 3$, et $N_{i,j} = 24$ si $R_{i,j} \geq 3; i \in \llbracket 1, \mathcal{L} \rrbracket; j \in \llbracket R_i, R_{i+1} \rrbracket$. Cette représentation est basée sur le schéma multi-échelles de LIU et collab. [2014] et dont l'intérêt est de limiter la sur-représentation du voisinage et éventuellement l'encodage d'information potentiellement bruitée. La Figure 4.7 montre un exemple d'un voisinage dense pour des macrostructures *circ* et *ring* avec $\mathcal{L} = 3$.

Le schéma de distribution des points voisins *ALLN* est largement utilisé pour une représentation multi-échelles dans différentes variantes LBP LIU et collab. [2017]. Cependant, vu sa limite principale qui est la sensibilité au bruit pour des macrostructures de grande taille MÄENPÄÄ et PIETIKÄINEN [2003], la plupart de ces méthodes se sont limitées à trois niveaux d'échelle et/ou ont combiné des mécanismes de tolérance au bruit LIAO et collab. [2007]; MÄENPÄÄ et PIETIKÄINEN [2003]; WANG et collab. [2013].

FIGURE 4.8 – Distribution creuse des points voisins *OneN*.

4.1.2.2 Voisinage creux : *OneN*

Le schémas *OneN* reflète une faible densité des points N_i dans le voisinage multi-échelles selon une information directionnelle exprimée par la valeur de θ . Dans ce travail, on a retenu les voisins répartis selon une valeur de déplacement angulaire $\theta = \pi/4$ suivant les quatre directions principales, à savoir, $0^\circ, 45^\circ, 90^\circ, 135^\circ$ pour un total de points $N_i = 8$ à chaque niveau d'échelle (c.f Figure 4.8).

On présume que huit points par niveau d'échelle, $\{N_i = 8\}_{i=1}^{\mathcal{L}}$, contiennent une information directionnelle de la macrostructure qui permet d'extraire des attributs informatifs. D'autre part, ce schéma était adopté dans plusieurs variantes multi-échelles LBP LIAO et collab. [2007]; LIU et collab. [2016b]; MÄENPÄÄ et PIETIKÄINEN [2003]; SIXTA [2011]; ŠULC et MATAS [2013]; WOLF et collab. [2008] pour les mêmes finalités d'optimisation.

4.2 Distribution d'intensité

Une fois la macrostructure extraite, une seconde étape du traitement est proposée dans notre approche, elle consiste en la représentation de l'information d'intensité contenue dans la macrostructure. Cette représentation est réalisée par l'application \mathcal{R} comme suit :

$$\begin{aligned} \mathcal{R} : \mathbb{R}^{N_i} &\rightarrow \mathbb{R}^S \\ \mathcal{X}_{R_i, N_i} &\mapsto \mathcal{R}(\mathcal{X}_{R_i, N_i}) \\ \mathcal{R}(\mathcal{X}_{R_i, N_i}) &= \mathcal{S}_i \end{aligned} \quad (4.3)$$

L'objectif de cette étape est de produire une représentation compacte, informative, et robuste à partir des valeurs d'intensité des points voisins dans chacun des niveaux d'échelle de la macrostructure. Cependant, il est nécessaire que ces représentations soient homogènes pour tous les niveaux d'échelle afin d'exprimer une information commune dont on peut encoder les variations dans la prochaine étape du traitement.

4.2.1 Mesures Statistiques

On propose une représentation statistique de la macrostructure. Cette représentation est essentiellement basée sur le résumé de la distribution de l'information d'intensité par des attributs intrinsèques calculés à partir des valeurs d'intensité des points voisins à chaque niveau de la macrostructure.

De ce fait, l'application \mathcal{R} réalise une opération de projection de l'espace d'intensité vers l'espace de distribution, où les valeurs d'intensité seront représentées par des mesures de statistique descriptive qui résument leur distribution.

Ce choix de représentation statistique est argumenté par les nombreux avantages offerts par les mesures statistique quant à la représentation de l'information dans la macrostructure :

- ▶ Les outils de la statistique descriptive sont des mesures statistiques simples à calculer, sans passer par une étape d'apprentissage pour déterminer les représentants dominants pour un large ensemble de données.
- ▶ Les mesures statistiques permettent de représenter la distribution de l'information d'intensité dans la macrostructure, et peut modéliser de façon simple et fiable les variations d'intensité qui caractérisent la surface de l'écorce.
- ▶ La statistique descriptive permet de *résumer* et de *représenter* la distribution d'un grand nombre de données par des représentants intrinsèques pertinents. Dans notre cas, cela permet de résumer la distribution des valeurs d'intensité des points voisins dans la macrostructure par un nombre réduit de représentants statistiques. Ainsi, l'information au sein de la macrostructure est compactée mais aussi préservée.
- ▶ Appliquer une description statistique sur un ensemble de points dans la macrostructure assure certaines propriétés d'invariance, notamment l'invariance aux changements locaux de l'orientation des structures de l'écorce (dus aux conditions naturelles et environnementales). En effet, une représentation dans l'espace de distribution est indépendante de l'emplacement spatial des points voisins et leur répartition le long de la structure du voisinage.

Nous définissons dans ce qui suit quelques mesures de la statistique descriptive. Par la suite, nous détaillons la représentation statistique de l'information dans la macrostructure via l'application \mathcal{R} .

4.2.1.1 Statistique descriptive

L'objectif de la statistique descriptive est de fournir une représentation et un résumé quantitatif d'un ensemble de données volumineux. Ceci est réalisable par le calcul de mesures statistiques à partir de cet ensemble de données. Ces mesures sont des représentants intrinsèques et informatifs de la distribution des données.

Les outils les plus communément utilisés pour résumer la distribution sont répartis en trois catégories : (1) les mesures de la tendance centrale et de la position, (2) les mesures de dispersion, et (3) les mesures qui décrivent la forme de la distribution.

Les mesures statistiques de la tendance centrale et de la position (*Central Tendency* (CT)) permettent de déterminer le point autour duquel les données d'un ensemble \mathcal{D} ont tendance à se

centrer. Ces mesures indiquent aussi la position dans l'espace de distribution. Les représentants les plus utilisés pour la centralité sont la moyenne, la médiane, le mode, et pour la position, les quantiles.

Les valeurs statistiques de la dispersion mesurent la variation de la distribution des données, et indiquent à quel degré les données sont étirées ou compactées par rapport à la moyenne.

Les mesures de la dispersion et l'étendue des données (*Dispersion and Range* (DR)), sont des valeurs réelles non négatives. Plus la valeur est grande, plus la dispersion et la variabilité des données est grande. Les valeurs statistiques de dispersion les plus fréquemment utilisées sont la variance, l'écart type, et l'étendue.

La description de la forme de la distribution des données peut être mesurée par les indices d'asymétrie et l'aplatissement.

De façon générale, soit un ensemble volumineux \mathcal{D} de N données discrètes défini comme suit :

$$\mathcal{D} = \{x_1, x_2, \dots, x_N\} \quad (4.4)$$

La description de la distribution de ces données pour chacune des catégories descriptives est présentée en détail dans ce qui suit :

La moyenne

C'est le représentant statistique le plus utilisé et le plus informatif pour décrire la centralité de la distribution des données. Sa valeur est donnée par la formule suivante :

$$Mean_{\mathcal{D}} = \frac{1}{N} \sum_{n=1}^N x_n \quad (4.5)$$

La distribution est centrée autour de la moyenne si la valeur de celle-ci est proche des valeurs des données x_n . Cependant, la moyenne est sensible aux valeurs aberrantes. Cela peut engendrer une information non représentative pour cet ensemble de données. Une solution consiste à tronquer ou éliminer un faible pourcentage de l'ensemble des données (un $p\%$ des valeurs extrêmes) et calculer la moyenne sur le reste. Cependant, cette solution néglige une partie de l'ensemble de données, ce qui peut conduire à une perte d'information. De ce fait, d'autres mesures de centralité sont utilisées, comme la médiane.

La médiane

La valeur médiane est la valeur centrale qui divise la distribution de l'ensemble de données en deux parties ayant un nombre égal de données. Ainsi, 50% de l'ensemble des données est inférieur à la médiane et l'autre 50% en est supérieur.

Formellement, l'ensemble \mathcal{D} est d'abord trié dans l'ordre croissant donnant : $\{x_1 = \min(\mathcal{D}), x_2, \dots, x_n = \max(\mathcal{D})\}$, ensuite, la valeur de la médiane est calculée comme suit :

$$Med_{\mathcal{D}} = \begin{cases} \frac{x_{(N/2)} + x_{(1+N/2)}}{2}, & \text{si } N \text{ est impair} \\ x_{((N+1)/2)}, & \text{si } N \text{ est pair} \end{cases} \quad (4.6)$$

La médiane est moins sensible aux valeurs extrêmes. Cependant, le tri de l'ensemble de données devient coûteux en temps de calcul et en stockage lorsque la taille des données est conséquente.

Le mode

Cette mesure représente la valeur la plus fréquente dans un ensemble de données. Le mode est généralement déterminé à partir de l'histogramme des fréquences des x_n valeurs de l'ensemble \mathcal{D} .

Les quantiles

Les quantiles sont les valeurs qui divisent l'ensemble de données en intervalles ayant la même taille (le même nombre de données dans chaque intervalle). Ces valeurs renseignent sur la position et sont calculées à partir de l'ensemble trié des données. Les divisions les plus communément utilisées sont les quartiles, les déciles, et les centiles.

- ▶ **Les quartiles** : ce sont les trois valeurs qui divisent l'ensemble de données triées en quatre intervalles, notées Q_1 , Q_2 , et Q_3 . Q_1 est le premier quartile. Il sépare les 25% des données inférieures à sa valeur et les 75% supérieures à sa valeur. Q_2 sépare les 50% inférieures des données. Autrement dit, Q_2 divise l'ensemble de données triées en deux, il est appelé aussi la médiane. Enfin, Q_3 , le troisième quartile dont 75% des données sont inférieures à sa valeur et 25% en sont supérieures. Différentes méthodes de calcul des quartiles est sont revues dans LANGFORD [2006].
- ▶ **Les déciles** : ce sont les neuf quantiles qui divisent l'ensemble de données en dix intervalles. Le premier décile sépare les premiers 10% des données inférieures du reste, et la valeur du neuvième décile est supérieure à 90% des données et inférieure à 10%.
- ▶ **Les centiles** : les centiles, quant à eux, divisent l'ensemble des données triées en cent quantiles. Le premier centile sépare les 1% des données inférieures à sa valeur, et le dernier centile, le 99^{ième}, a la valeur inférieure à 1% des données ayant de grandes valeurs dans l'ensemble.

La variance

La variance mesure la variabilité de la distribution des données autour de la moyenne. Sa valeur est donnée par :

$$\text{Var}_{\mathcal{D}} = \frac{1}{N} \sum_{n=1}^N (x_n - \text{Mean}_{\mathcal{D}})^2 \quad (4.7)$$

où $\text{Mean}_{\mathcal{D}}$ est la moyenne de l'ensemble des données \mathcal{D} .

L'écart type

C'est la racine au carré de la variation. L'écart type (*standard deviation*) est obtenu par :

$$Std_{\mathcal{D}} = \sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - Mean_{\mathcal{D}})^2} \quad (4.8)$$

Les valeurs de l'écart type ont le même ordre de grandeur que les valeurs des données.

La valeur minimale

C'est la valeur la plus petite dans tout l'ensemble des données :

$$Min_{\mathcal{D}} = \min(\{x_n\}_{n=1}^N) \quad (4.9)$$

La valeur maximale

C'est la valeur la plus grande dans tout l'ensemble des données :

$$Max_{\mathcal{D}} = \max(\{x_n\}_{n=1}^N) \quad (4.10)$$

L'étendue

Étant $Min_{\mathcal{D}}$, et $Max_{\mathcal{D}}$, l'étendue est la différence entre les valeurs extrêmes. Sa valeur est calculée par :

$$Etendue_{\mathcal{D}} = Max_{\mathcal{D}} - Min_{\mathcal{D}} \quad (4.11)$$

L'asymétrie

Dite *skeweness*, l'asymétrie interprète la répartition des données autour de la moyenne. La valeur de l'indice d'asymétrie est calculée par la formule suivante :

$$\gamma_{1,\mathcal{D}} = \frac{\frac{1}{N} \sum_{n=1}^N (x_n - Mean_{\mathcal{D}})^3}{Std_{\mathcal{D}}^3} \quad (4.12)$$

où $Mean_{\mathcal{D}}$ est la moyenne de la distribution et $Std_{\mathcal{D}}$ est son écart-type.

Les valeurs positives de $\gamma_{1,\mathcal{D}}$ indiquent que la distribution des données est décalée à gauche de la moyenne, et inversement pour les valeurs négatives.

L'aplatissement

Aussi appelé l'acuité (*kurtosis*), l'aplatissement indique si la distribution des données est plus ou moins sujette aux valeurs aberrantes. Sa valeur est donnée par :

$$\gamma_{2,\mathcal{D}} = \frac{\frac{1}{N} \sum_{n=1}^N (x_n - Mean_{\mathcal{D}})^4}{Std_{\mathcal{D}}^4} \quad (4.13)$$

où $Mean_{\mathcal{D}}$ est la moyenne de la distribution et $Std_{\mathcal{D}}$ est son écart-type.

4.2.1.2 Moments statistiques

Par abus de langage, on retrouve d'autres appellations dans la littérature pour exprimer les mesures statistiques relatives à la distribution d'information. Il s'agit des moments statistiques. Les moments sont utilisés pour approximer la distribution des données d'un ensemble fini, et se calculent par ordre de grandeurs o . Cependant, c'est le même principe de calcul qui s'applique pour les moments statistiques d'ordre réduit que celui pour les mesures de la statistique descriptive. En effet, on retrouve :

- moment ordinaire d'ordre zéro étant égal à 1.
- moment ordinaire d'ordre un, i.e., la moyenne $Mean_{\mathcal{D}}$.
- moment centré d'ordre deux, i.e., la variance $Var_{\mathcal{D}}$.
- moment centré réduit d'ordre trois, i.e., l'asymétrie $\gamma_{1,\mathcal{D}}$.
- moment centré réduit d'ordre quatre, i.e., l'aplatissement $\gamma_{2,\mathcal{D}}$.

Les moments d'ordre $o \geq 4$ sont appelés les moments d'ordre très élevé (*high order moments*). Ils se caractérisent par la complexité élevée de calcul, et la sensibilité au bruit. De ce fait, et en raison de la complexité de calcul et la faible performance, la plupart des approximations incluent seulement les moments d'ordre zéro à quatre.

4.2.1.3 Autres mesures statistiques

En plus des outils de la statistique descriptive qui décrivent la distribution des données au sein d'un seul ensemble, il existe des mesures qui représentent la relation de dépendance statistique entre deux ensembles de données ayant la même taille. L'une des mesures les plus utilisées est celle du coefficient de corrélation. Il est défini entre les éléments de deux ensembles ayant le même nombre de données par la formule suivante :

$$\rho(\mathcal{D}, \mathcal{B}) = \frac{1}{N} \sum_{n=1}^N \left(\frac{x_n - Mean_{\mathcal{D}}}{Std_{\mathcal{D}}} \right) \left(\frac{x_n - Mean_{\mathcal{B}}}{Std_{\mathcal{B}}} \right) \quad (4.14)$$

$\rho(\mathcal{D}, \mathcal{B})$ prend des valeurs dans l'intervalle $[-1, 1]$. La valeur 0 indique l'absence de relation linéaire entre les deux ensembles. Il prend les valeurs 1 ou -1 si tous les points sont alignés sur une droite.

4.2.2 Représentation statistique

Chacune des mesures statistiques abordées représente un aspect particulier de la distribution des données. En appliquant ces outils à notre étude, différentes représentations de la distribution d'intensité peuvent être dérivées pour une macrostructure donnée. Cependant, la principale motivation du choix des mesures statistiques se focalise sur l'intérêt de produire une représentation *similaire* de la texture des écorces appartenant à la même espèce végétale, et non pas une représentation *identique*. Autrement dit, on souhaite extraire des attributs intrinsèques pouvant générer des descripteurs similaires pour des images d'écorce de la même classe, sans pour autant biaiser la description et limiter sa généralisation pour la description des espèces avec une grande variabilité d'apparence de leurs surface. En effet, procéder par une approximation fiable de la distribution d'intensité pose une contrainte d'identité sur les images d'espèce, or ceci ne peut être

applicable pour des objets naturels. D'autant plus, une caractérisation *similaire* modélise le cas d'application effectuée par les botanistes pour déterminer l'espèce d'un nouvel échantillon : ils cherchent les caractéristiques visuelles les plus similaires pour identifier une espèce donnée.

Dans cette perspective, on procède par une approximation de la distribution d'intensité dans la macrostructure par des statistiques de bas ordre, car limité par les contraintes suivantes :

- ▶ Une approximation de la distribution d'intensité par l'ensemble des moments de l'ordre $o = 0$ au moment d'ordre $o \rightarrow \infty$ permet la reconstruction de l'ensemble de données, à l'encontre du cadre de notre application, où on ne cherche pas une approximation fiable de la distribution.
- ▶ La complexité de calcul des moments d'ordre élevé est proportionnel à leur ordre.
- ▶ Les moments statistiques d'ordre élevé ont tendance à amplifier les variations, et donc à encoder une information instable et bruitées, ce qui les rend moins informatifs et donc moins performants NGUYEN et collab. [2016].

Par conséquent, la description de la distribution d'intensité se restreint aux moments d'ordre un, i.e, la moyenne (Mean), et le moment centré d'ordre deux, i.e, la variance (Var). De plus, on procède à la représentation de l'information de distribution de ces régions par les deux valeurs statistiques de minimum (Min) est maximum (Max) respectivement. De plus, compte tenu des conditions d'acquisition des images d'écorce non-contrôlées, une autre mesure est aussi considérée : la médiane (Med), qui est plus stable que la moyenne, et est moins sensible au bruit du en majeure partie aux changements de luminance.

Formellement, on définit un ensemble de valeurs statistiques \mathcal{M} qui modélisent une représentation spécifique de l'apparence de la surface d'écorce comme suit : $\mathcal{M} = \{Mean, Var, Min, Max, Med\}$

L'ensemble \mathcal{M} constitue aussi l'espace de projection de l'application \mathcal{R} selon la formule suivante :

$$\begin{aligned} \mathcal{R} : \mathbb{R}^{N_i} &\rightarrow \mathbb{R}^{|\mathcal{M}|} \\ \mathcal{X}_{R_i, N_i} &\mapsto \mathcal{R}(\mathcal{X}_{R_i, N_i}) \\ \mathcal{R}(\mathcal{X}_{R_i, N_i}) &= \mathcal{S}_i \\ \mathcal{S}_i &= [\mathcal{S}_{i, m}]^T, \quad m \in \mathcal{M} \end{aligned} \tag{4.15}$$

où :

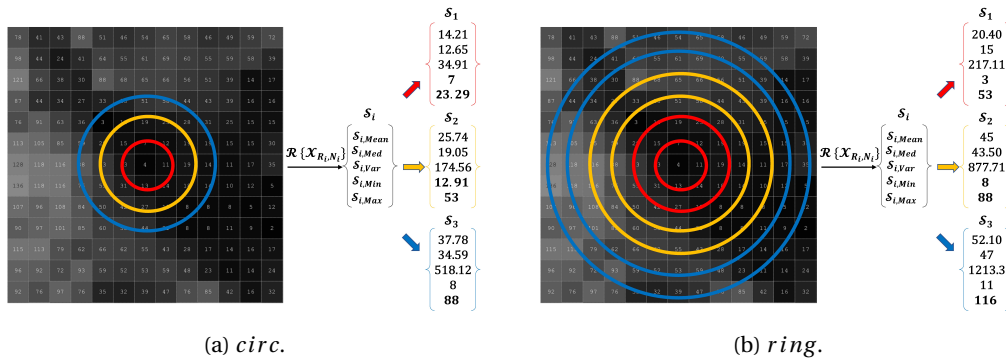
- $|\mathcal{M}|$ symbolise la cardinalité de l'ensemble \mathcal{M} , et donne aussi la dimension du nouvel espace de représentation. Dans cette étude, elle est égale à 5.
- \mathcal{S}_i est un vecteur statistique donné par : $\mathcal{S}_i = [Mean_i, Var_i, Min_i, Max_i, Med_i]^T$.

Chacune des composantes statistiques du vecteur $[\mathcal{S}_{i, m}]^T$ est calculée par les formules suivantes :

$$Mean_i = \frac{1}{N_i} \sum_{n=1}^{N_i} x_n \tag{4.16}$$

$$Var_i = \frac{1}{N_i} \sum_{n=1}^{N_i} (x_n - Mean_i)^2 \tag{4.17}$$

$$Min_i = \min(\{x_n\}_{n=1}^{N_i}) \tag{4.18}$$

FIGURE 4.9 – Exemple de description statistique avec $\mathcal{L} = 3$.

$$Max_i = \max(\{x_n\}_{n=1}^{N_i}) \quad (4.19)$$

$$Med_i = \frac{x_{(N_i/2)} + x_{(1+N_i/2)}}{2} \quad (4.20)$$

L'application \mathcal{R} est réalisée pour chacun des niveaux de l'espace d'échelle, i.e., les vecteurs d'intensité $\{\mathcal{X}_{R_i, N_i}\}_{i \in \llbracket 1, \mathcal{L} \rrbracket}$, ce qui résulte en un ensemble de vecteurs statistiques $\{\mathcal{S}_{i, m}, m \in \mathcal{M}\}_{\llbracket 1, \mathcal{L} \rrbracket}$ (c.f Figure 4.9).

4.3 Codes macro-motifs

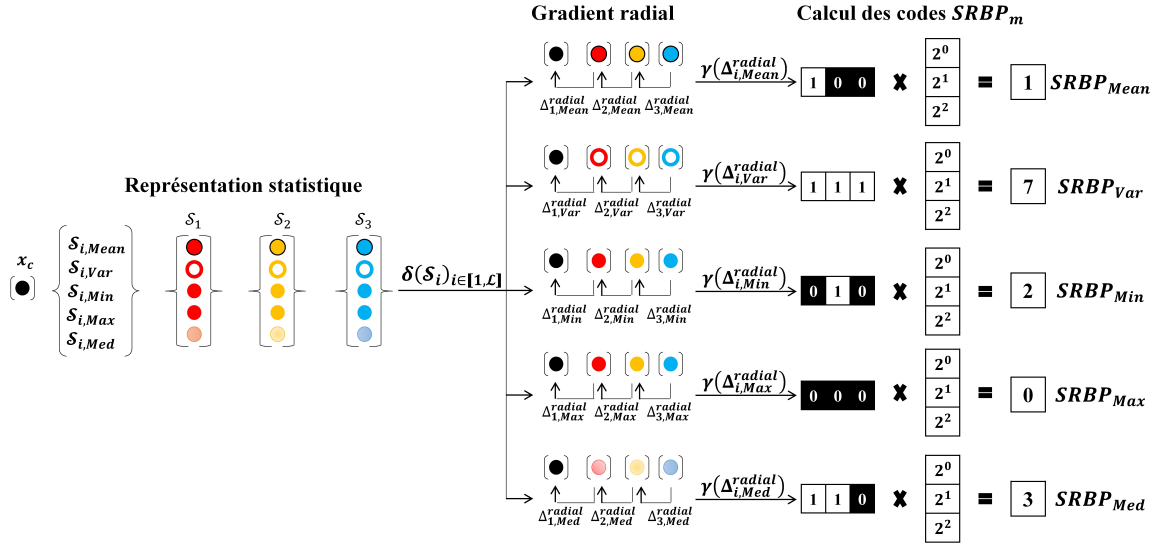
Suite à la projection dans l'espace de distribution, la macrostructure est représentée par un ensemble de vecteurs statistiques de taille réduite et uniforme. Chaque niveau d'échelle de la macrostructure, i.e., un cercle ou un anneau, est représenté par un vecteur statistique. Ceci produit une représentation homogène autour du pixel central x_c . Cette homogénéité de la représentation de l'information pour les niveaux d'échelle offre la possibilité d'encoder la relation entre ces niveaux d'échelle selon différents schémas d'encodage afin de générer un code *macro-motif* caractéristique de cette macrostructure.

Dans le contexte de notre travail, on modélise l'étape d'encodage de l'information par l'application \mathcal{C} comme suit :

$$\begin{aligned} \mathcal{C} : \mathbb{R}^{\mathcal{L}} &\rightarrow \llbracket 0, 2^{\mathcal{L}} - 1 \rrbracket \\ \{\mathcal{S}_i\} &\mapsto \mathcal{C}(\{\mathcal{S}_i\}) \\ \mathcal{C}(\{\mathcal{S}_i\}) &= \text{macro-motif}, \quad |i \in \llbracket 1, \mathcal{L} \rrbracket \end{aligned} \quad (4.21)$$

Il s'agit d'une projection de l'ensemble des vecteurs statistiques \mathcal{S}_i dans l'espace code, où, selon un encodage défini, on construit une chaîne binaire de taille \mathcal{L} . Cette chaîne binaire est ensuite convertie en décimal, donnant un code macro-motif. Ce code macro-motif représente un pixel donné de l'image et caractérise la structure de l'écorce autour de lui.

Le code macro-motif est un encodage des variations entre les niveaux d'échelle de la macrostructure. Cette variation est quantifiée par la valeur d'un gradient selon une direction donnée Δ_i^{dir} . En fonction de cette direction, une information particulière sur la distribution d'information dans la macrostructure est encodée. On introduit deux types différents de gradient, nommés

FIGURE 4.10 – Calcul des codes macro-motifs radiaux $SRBP_m$.

Δ_i^{radial} pour un gradient radial, et Δ_i^{depth} pour un gradient en profondeur. Par conséquent, deux type de code macro-motifs sont présentés dans ce qui suit.

4.3.1 Macro-motif binaire statistique radial

Le code macro-motif binaire statistique radial repose sur le calcul du gradient radial. Ce gradient caractérise l'information inter-échelles en captant les variations entre deux niveaux consécutifs de l'espace d'échelle. Ceci permet aussi de représenter la distribution graduelle de l'information d'un niveau d'échelle à un niveaux plus supérieur dans la macrostructure. On définit le gradient radial par :

$$\Delta_i^{radial} = [S_{i,m}]^T - [S_{i-1,m}]^T; [S_{1,m}]^T - [x_c]; i \in [1, \mathcal{L}]; m \in \mathcal{M} \quad (4.22)$$

tel que x_c est le pixel central. La Figure 4.10 montre un schéma de calcul du gradient radial.

On dénomme un code macro-motif radial par $SRBP_m$ (*Statistical Radial Binary Pattern*) pour une représentation statistique donnée $m \in \mathcal{M}$. Un code $SRBP_m$ est calculé par :

$$SRBP_m = \sum_{i=1}^{\mathcal{L}} f(S_{i,m} - S_{i-1,m}) 2^{i-1}; \quad f(S_{i,m} - S_{i-1,m}) = \begin{cases} 1 & \text{si } S_{i,m} \geq S_{i-1,m} \\ 0 & \text{sinon.} \end{cases} \quad (4.23)$$

où $S_{0,m} = x_c | m \in \mathcal{M}$.

Dans ce cas, le seuil de binarisation adaptatif pour chaque niveau de l'échelle $S_{i,m}$ est le niveau d'échelle précédent $S_{i-1,m}$. La Figure 4.10 montre un exemple de calcul des codes macro-motifs $SRBP_m$ pour chacune des représentations statistiques de l'image d'écorce avec une macrostructure de trois échelles $\mathcal{L} = 3$.

4.3.2 Macro-motif binaire statistique profond

Pour le calcul du code macro-motif binaire statistique profond on utilise le gradient en profondeur Δ_i^{depth} . Ce gradient caractérise la relation entre les niveaux d'échelle en les comparant

à une représentation moyenne de la distribution de l'information dans la macrostructure. Ainsi on représente une information globale sur la macrostructure dont les variations entre ses échelles sont mesurées par rapport à un prototype, dénoté μ . Le calcul du gradient Δ_i^{depth} est donné par :

$$\Delta_i^{depth} = [\mathcal{S}_{i,m}]^T - [\mu_m]^T; i \in \llbracket 1, \mathcal{L} \rrbracket; m \in \mathcal{M} \quad (4.24)$$

où $[\mu_m]_{m \in \mathcal{M}}^T$ est un prototype statistique intrinsèque qui représente une valeur moyenne calculée à partir des vecteurs $\{[\mathcal{S}_{i,m}]_{i \in \llbracket 1, \mathcal{L} \rrbracket; m \in \mathcal{M}}^T$.

Par ailleurs, deux types de prototypes μ sont définis selon le niveau de représentation :

- ▶ Un prototype intrinsèque global propre au niveau de l'échelle i , dénoté par μ_i^{scale} .
- ▶ Un prototype intrinsèque local propre au macro-motif, dénoté par $\mu^{pattern}$.

4.3.2.1 Prototype intrinsèque global

Soit les représentations $\{[\mathcal{S}_{i,m}]_{i \in \llbracket 1, \mathcal{L} \rrbracket; m \in \mathcal{M}}^T$ d'une macrostructure autour d'un pixel donné dans l'image \mathcal{I} , le prototype intrinsèque global μ^{scale} est un vecteur moyen 2D de taille $\mathcal{L} \times |\mathcal{M}|$ représentatif de tous les vecteurs $[\mathcal{S}_{i,m}]^T$ pour toutes les macrostructures considérées dans l'image \mathcal{I} . Formellement, pour chaque $i \in \llbracket 1, \mathcal{L} \rrbracket$ et chaque $m \in \{Mean, Var, Min, Max, Med\}$, μ^{scale} est donné par :

$$\mu^{scale} = \begin{bmatrix} \mu_{1,Mean}^{scale} & \mu_{2,Mean}^{scale} & \cdots & \mu_{\mathcal{L},Mean}^{scale} \\ \mu_{1,Var}^{scale} & \mu_{2,Var}^{scale} & \cdots & \mu_{\mathcal{L},Var}^{scale} \\ \mu_{1,Min}^{scale} & \mu_{2,Min}^{scale} & \cdots & \mu_{\mathcal{L},Min}^{scale} \\ \mu_{1,Max}^{scale} & \mu_{2,Max}^{scale} & \cdots & \mu_{\mathcal{L},Max}^{scale} \\ \mu_{1,Med}^{scale} & \mu_{2,Med}^{scale} & \cdots & \mu_{\mathcal{L},Med}^{scale} \end{bmatrix} \quad (4.25)$$

Autrement dit, soit l'image \mathcal{I} , pour une représentation statistique donnée m , la projection dans l'espace de distribution d'un niveau d'échelle spécifique i génère une image $\mathcal{I}_{i,m}$. La moyenne globale de l'image $\mathcal{I}_{i,m}$ produit la valeur $\mu_{i,m}^{scale}$ du prototype statistique intrinsèque global. Formellement :

$$\mu_{i,m}^{scale} = \frac{1}{H \times W} \sum_{i=1}^{H \times W} \mathcal{I}_{i,m} \quad (4.26)$$

où :

H = hauteur de l'image \mathcal{I} – hauteur de la macrostructure.

W = largeur de l'image \mathcal{I} – largeur de la macrostructure.

La Figure 4.11 montre un exemple de calcul du prototype global pour une macrostructure de $\mathcal{L} = 3$ niveaux. Dans cet exemple μ^{scale} est donné par :

$$\mu^{scale} = \begin{bmatrix} \mu_{1,Mean}^{scale} & \mu_{2,Mean}^{scale} & \mu_{3,Mean}^{scale} \\ \mu_{1,Var}^{scale} & \mu_{2,Var}^{scale} & \mu_{3,Var}^{scale} \\ \mu_{1,Min}^{scale} & \mu_{2,Min}^{scale} & \mu_{3,Min}^{scale} \\ \mu_{1,Max}^{scale} & \mu_{2,Max}^{scale} & \mu_{3,Max}^{scale} \\ \mu_{1,Med}^{scale} & \mu_{2,Med}^{scale} & \mu_{3,Med}^{scale} \end{bmatrix} \quad (4.27)$$

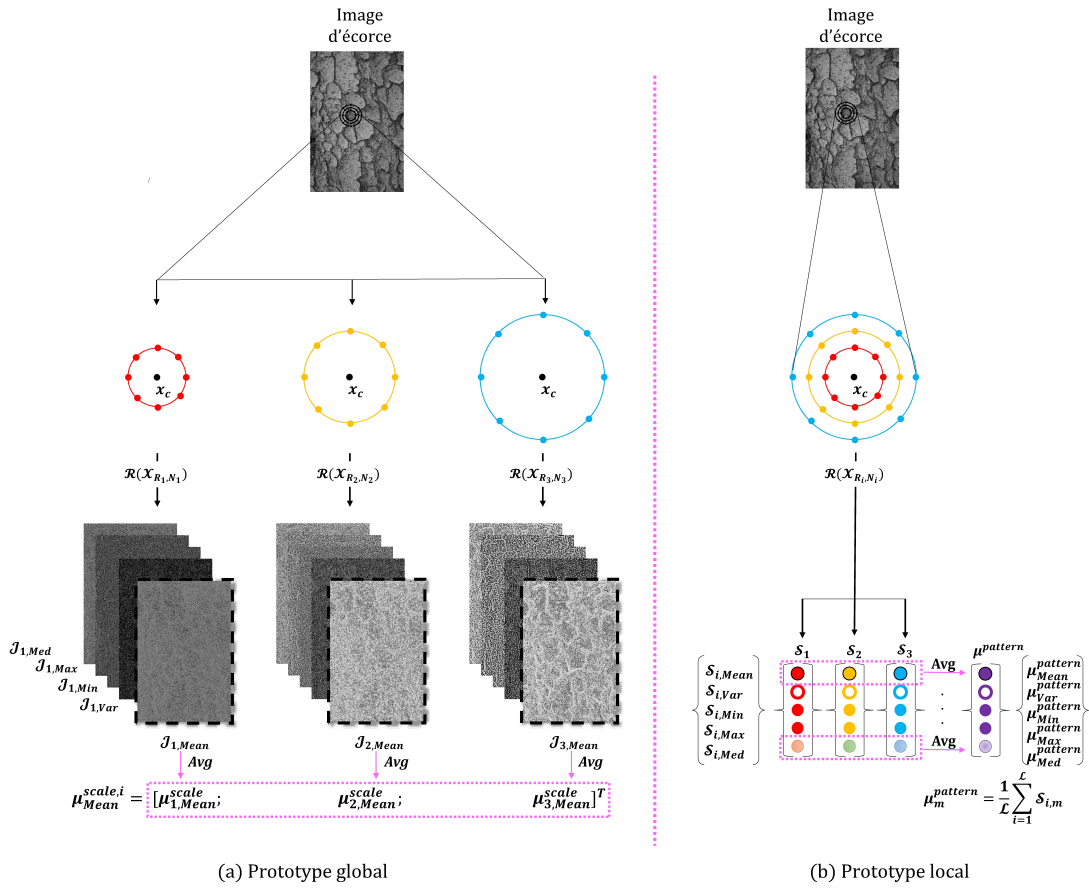


FIGURE 4.11 – Calcul des prototypes statistiques intrinsèques.

et les vecteurs gradients $\{\Delta_i\}_{i \in [1,3]}^{depth}$ sont ainsi donnés par :

$$\Delta_1^{depth} = \begin{bmatrix} \mathcal{S}_{1,Mean} - \mu_{1,Mean}^{scale} \\ \mathcal{S}_{1,Var} - \mu_{1,Var}^{scale} \\ \mathcal{S}_{1,Min} - \mu_{1,Min}^{scale} \\ \mathcal{S}_{1,Max} - \mu_{1,Max}^{scale} \\ \mathcal{S}_{1,Med} - \mu_{1,Med}^{scale} \end{bmatrix}; \Delta_2^{depth} = \begin{bmatrix} \mathcal{S}_{2,Mean} - \mu_{2,Mean}^{scale} \\ \mathcal{S}_{2,Var} - \mu_{2,Var}^{scale} \\ \mathcal{S}_{2,Min} - \mu_{2,Min}^{scale} \\ \mathcal{S}_{2,Max} - \mu_{2,Max}^{scale} \\ \mathcal{S}_{2,Med} - \mu_{2,Med}^{scale} \end{bmatrix}; \Delta_3^{depth} = \begin{bmatrix} \mathcal{S}_{3,Mean} - \mu_{3,Mean}^{scale} \\ \mathcal{S}_{3,Var} - \mu_{3,Var}^{scale} \\ \mathcal{S}_{3,Min} - \mu_{3,Min}^{scale} \\ \mathcal{S}_{3,Max} - \mu_{3,Max}^{scale} \\ \mathcal{S}_{3,Med} - \mu_{3,Med}^{scale} \end{bmatrix} \quad (4.28)$$

Le prototype intrinsèque globale μ^{scale} représente une information globale adaptative par niveau de résolution spatiale. En effet, on calcule un vecteur moyenne par niveau de résolution spatiale pour toutes les macrostructures de l'image de texture d'écorce. μ^{scale} mesure la centralité de la distribution de l'information d'intensité par niveau d'échelle, et son utilisation pour mesurer la relation entre les niveaux d'échelle reflète la variation de la distribution d'information par résolution sur toute l'image de la texture d'écorce.

4.3.2.2 Prototype intrinsèque local

Différemment au prototype global, le prototype intrinsèque local $\mu^{pattern}$ est un vecteur moyen de taille $|\mathcal{M}|$ calculé à partir de toutes les représentations statistiques $\{\mathcal{S}_{i,m}\}_{i \in [1,\mathcal{L}]}$ pour une macrostructure donnée. Plus spécifiquement, pour $m \in \{Mean, Var, Min, Max, Med\}$, $\mu^{pattern}$ est

donné par :

$$\mu^{pattern} = \begin{bmatrix} \mu_{Mean}^{pattern} \\ \mu_{Var}^{pattern} \\ \mu_{Min}^{pattern} \\ \mu_{Max}^{pattern} \\ \mu_{Med}^{pattern} \end{bmatrix} \quad (4.29)$$

tel que pour une mesure statistique spécifique m , $\mu_m^{pattern}$ est une moyenne locale donnée par :

$$\mu_m^{pattern} = \frac{1}{\mathcal{L}} \sum_{i=1}^{\mathcal{L}} S_{i,m} \quad (4.30)$$

La Figure 4.11 montre un exemple de calcul du prototype local pour une macrostructure de $\mathcal{L} = 3$ niveaux. Dans cet exemple les valeurs de $\mu_m^{pattern}$; $m \in \mathcal{M}$ sont données par :

$$\begin{aligned} \mu_{Mean}^{pattern} &= (\sum_{i=1}^3 S_{i,Mean})/3 \\ \mu_{Var}^{pattern} &= (\sum_{i=1}^3 S_{i,Var})/3 \\ \mu_{Min}^{pattern} &= (\sum_{i=1}^3 S_{i,Min})/3 \\ \mu_{Max}^{pattern} &= (\sum_{i=1}^3 S_{i,Max})/3 \\ \mu_{Med}^{pattern} &= (\sum_{i=1}^3 S_{i,Med})/3 \end{aligned} \quad (4.31)$$

et les vecteurs gradients $\{\Delta_i\}_{i \in [1,3]}^{depth}$ sont donnés par :

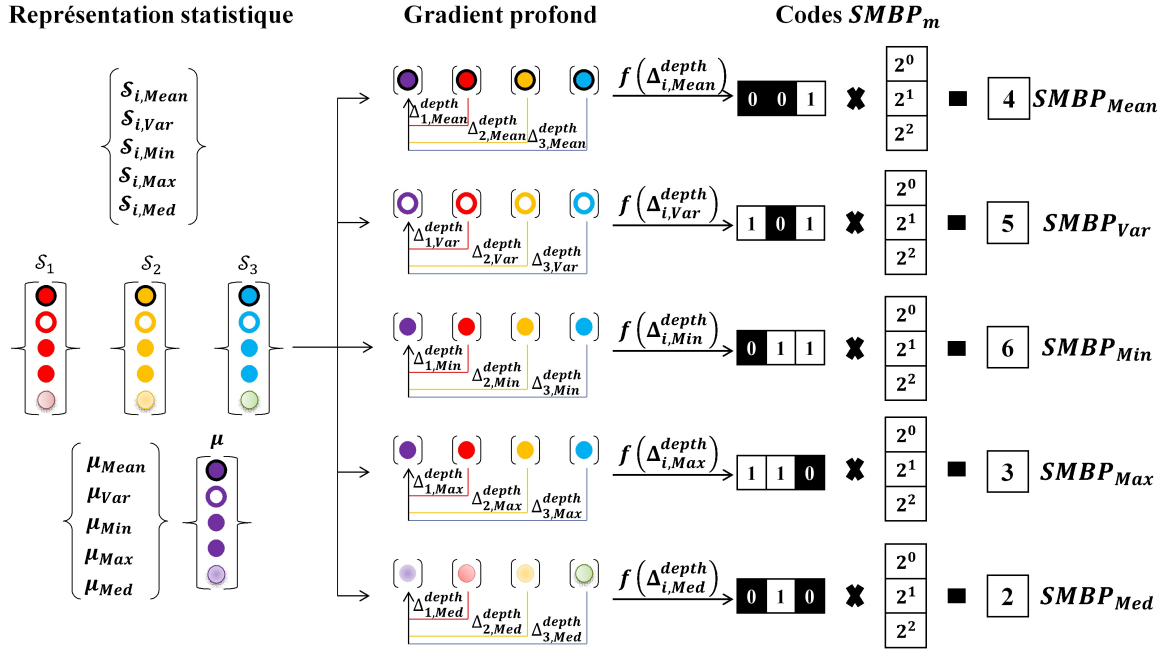
$$\Delta_1^{depth} = \begin{bmatrix} S_{1,Mean} - \mu_{Mean}^{pattern} \\ S_{1,Var} - \mu_{Var}^{pattern} \\ S_{1,Min} - \mu_{Min}^{pattern} \\ S_{1,Max} - \mu_{Max}^{pattern} \\ S_{1,Med} - \mu_{Med}^{pattern} \end{bmatrix}; \Delta_2^{depth} = \begin{bmatrix} S_{2,Mean} - \mu_{Mean}^{pattern} \\ S_{2,Var} - \mu_{Var}^{pattern} \\ S_{2,Min} - \mu_{Min}^{pattern} \\ S_{2,Max} - \mu_{Max}^{pattern} \\ S_{2,Med} - \mu_{Med}^{pattern} \end{bmatrix}; \Delta_3^{depth} = \begin{bmatrix} S_{3,Mean} - \mu_{Mean}^{pattern} \\ S_{3,Var} - \mu_{Var}^{pattern} \\ S_{3,Min} - \mu_{Min}^{pattern} \\ S_{3,Max} - \mu_{Max}^{pattern} \\ S_{3,Med} - \mu_{Med}^{pattern} \end{bmatrix} \quad (4.32)$$

Le prototype intrinsèque locale $\mu^{pattern}$ représente une information locale adaptative propre à chacune des macrostructures de l'image. En effet, $\mu^{pattern}$ est une moyenne statistique de la macrostructure calculée à partir de l'information statistique de ses niveaux d'échelle. Les valeurs de $\mu^{pattern}$ diffèrent d'une macrostructure à une autre, et son utilisation dans le calcul du gradient mesure la variation de la distribution d'information d'intensité dans la macrostructure.

De manière similaire aux codes macro-motifs radiaux, on dénote par $SMBP_m$ (*Statistical Macro Binary Pattern*) les codes macro-motifs profonds. Pour une représentation statistique spécifique $m \in \mathcal{M}$, un code $SMBP_m$ est donné par :

$$SMBP_m = \sum_{i=1}^{\mathcal{L}} f(S_{i,m} - \mu_m) 2^{i-1}; \quad f(S_{i,m} - \mu_m) = \begin{cases} 1 & \text{si } S_{i,m} \geq \mu_m \\ 0 & \text{sinon.} \end{cases} \quad (4.33)$$

Le seuil de binarisation adaptatif pour un code macro-motif profond est le prototype intrinsèque μ , tel qu'il peut être le prototype global μ^{scale} , ou bien le prototype local $\mu^{pattern}$. Un exemple de calcul des codes $SMBP_m$ avec un prototype local $\mu^{pattern}$ pour une macrostructure de $\mathcal{L} = 3$ niveaux est illustré dans la Figure 4.12.


 FIGURE 4.12 – Calcul des codes macro-motifs profonds $SMBP_m$.

4.4 Descripteurs de la texture d'écorce

L'objectif de cette thèse est de proposer de nouveaux descripteurs spécifiques pour la description de la structure d'écorce pour l'identification des espèces végétales. Les descripteurs introduits encodent communément l'information contenue dans la macrostructure en produisant un seul code macro-motif qui caractérise une grande région à des niveaux d'échelle élevés. Cependant l'encodage utilisé pour représenter l'information entre les niveaux d'échelle de la macrostructure est différents pour chacun des descripteurs proposés.

Afin de construire le descripteur final d'une image d'écorce, on procède par une représentation globale de l'image de texture de l'écorce, i.e., un code macro-motif est calculé pour chaque pixel de l'image en entrée. Ainsi, on obtient une image de codes statistiques qui caractérisent l'information structurale de l'image d'écorce. La distribution de ces codes macro-motifs est représentée par un histogramme de $2^{\mathcal{L}}$ bins où chaque code macro-motif vote dans son bin correspondant. La Figure 4.13 montre un exemple de construction des histogrammes $SRPB_m$ et $SMBP_m$, utilisant un prototype local $\mu^{pattern}$, pour une macrostructure de $\mathcal{L} = 3$.

4.4.1 SRBP : *Statistical Radial Binary Pattern*

Pour une image \mathcal{I} , le calcul de l'histogramme des codes radiaux $SRBP_m$ pour une représentation statistique $m \in \mathcal{M}$ est donné par :

$$SRBP_m[x_c] = \sum_{h=1}^H \sum_{w=1}^W SRBP_m(h, w) = c \quad | \quad c = 0, 1, 2, \dots, 2^{\mathcal{L}} - 1 \quad (4.34)$$

où

H = hauteur de l'image \mathcal{I} – hauteur de la macrostructure.

W = largeur de l'image \mathcal{I} – largeur de la macrostructure.

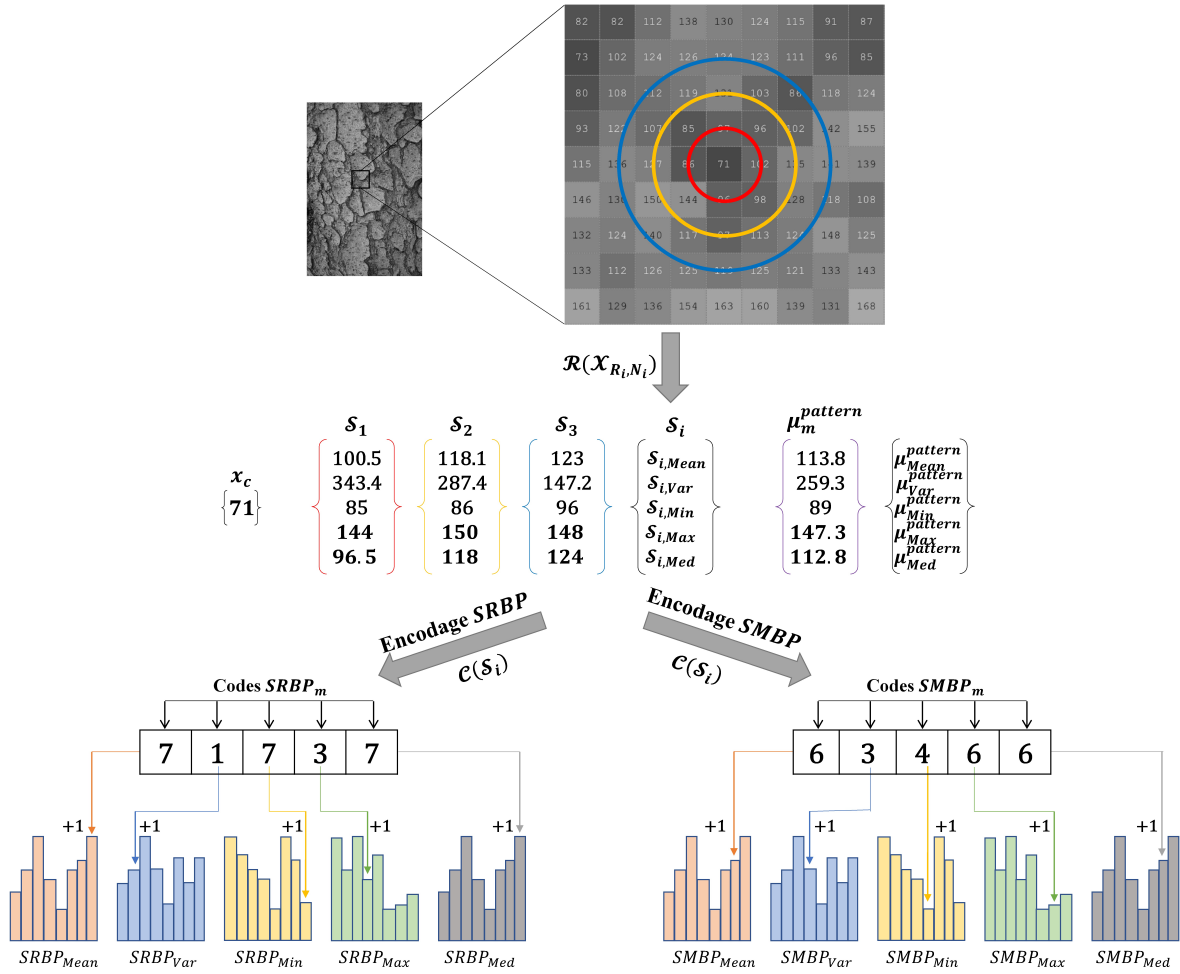


FIGURE 4.13 – Calcul des histogrammes des descripteurs $SRBP_m$ et $SMBP_m$. Exemples avec une macro-structure de $\mathcal{L} = 3$ niveaux d'échelle.

4.4.2 SMBP : *Statistical Macro Binary Pattern*

Suivant le même processus de représentation globale de l'image \mathcal{I} , le calcul de l'histogramme $SMBP_m$ des macro-motifs profonds pour une représentation statistique m est donné par :

$$SMBP_m[x_c] = \sum_{h=1}^H \sum_{w=1}^W SMBP_m(h, w) = c \quad | \quad c = 0, 1, 2, \dots, 2^{\mathcal{L}} - 1 \quad (4.35)$$

où

H = hauteur de l'image \mathcal{I} – hauteur de la macrostructure.

W = largeur de l'image \mathcal{I} – largeur de la macrostructure.

4.5 Propriétés des codes macro-motifs

Les descripteurs de texture proposés dans ce chapitre offrent différentes représentations de la structure d'écorce. Tout de même, ces représentations se focalisent toutes sur l'apport d'une information spécifique pour cerner les diverses apparences de la surface d'écorce.

À partir de la chaîne de traitement de la macrostructure, on compte les propriétés suivantes des descripteurs de texture d'écorce des macro-motifs binaires :

- ▶ L'approche proposée pour la description de la texture d'écorce est une généralisation de LBP et MResLBP. À l'encontre du schéma d'encodage de ces approches qui génère un code LBP par niveau de résolution et les combine, notre approche encode un voisinage multi-échelles par un seul code macro-motif qui reflète le mieux la relation entre les niveaux d'échelle de la macrostructure.
- ▶ La conception de l'espace d'échelle des descripteurs macro-motifs binaires exploite différentes distributions des niveaux de résolution et des points voisins à partir de larges supports spatiaux afin d'extraire des attributs informatifs et pertinents.
- ▶ Les descripteurs macro-motifs binaires encodent l'information à partir de l'espace de distribution, en calculant des attributs statistiques intrinsèques, qui représentent et résument la distribution d'intensité dans une large macrostructure.
- ▶ La plupart des variantes LBP ont l'inconvénient d'être sensibles aux variations de la rotation locale, d'où l'utilisation intensive du mapping *riu2* pour remédier à ce problème. Nos descripteurs de texture sont par définition invariants aux changements de la rotation locale. En effet, l'information retenue dans la macrostructure est une information statistique calculée à partir des valeurs d'intensité des points voisins indépendamment de leur emplacement spatial, ce qui assure cette invariance à la rotation locale, et améliore la robustesse du descripteur face à ce problème.
- ▶ Le principe de calcul du gradient est essentiellement basé sur un calcul de variation dans la distribution d'intensité par rapport à un prototype intrinsèque adaptatif. Pour les codes macro-motif radiaux, le prototype étant la distribution de l'échelle adjacente dans la macrostructure. Le code produit caractérise l'évaluation de la distribution d'intensité de façon graduelle dans la macrostructure. Pour les codes macro-motif profond, le prototype étant une moyenne statistique de la distribution d'intensité. Le code généré pour les différentes représentations de cette moyenne caractérise l'évaluation globale de la distribution d'intensité dans la macrostructure. Dans le cas des deux codes, procéder ainsi contribue à atténuer l'effet du bruit (régularisé par l'utilisation du prototype), et aussi de ressortir une information sur la relation entre les niveaux d'échelle dans la macrostructure.
- ▶ L'information encodée par nos descripteurs est uniquement basée sur un encodage binaire du signe du gradient, sans pour autant ajouter l'information complémentaire de la magnitude ou l'intensité du pixel central. Toutefois, la plupart des variantes LBP les plus performantes attribuent leurs succès à la combinaison d'information complémentaires, par représentation jointe le plus souvent, comme CLBP-S/M/C [GUO et collab. \[2010a\]](#), CLBC-S/M/C [ZHAO et collab. \[2012b\]](#), BRINT-CS-CM [LIU et collab. \[2014\]](#), et MC-SBP2 [NGUYEN et collab. \[2016\]](#). Dans notre approche, la complémentarité d'information qui contribuent à une description efficace de la texture d'écorce est réalisée par les différentes représentations statistiques de la distribution d'information dans la macrostructure.
- ▶ Les descripteurs statistiques proposés peuvent facilement être combinés pour ajouter une complémentarité d'information et améliorer la représentativité des structures d'écorce. En effet, les descripteurs statistiques sont uniformes en terme d'information encodée, i.e., la texture, le processus d'encodage, i.e., encodage de la macrostructure, et la dimension, i.e.,

tous de dimension $2^{\mathcal{L}}$. Ceci permet une fusion tardive par simple concaténation dans le but d'améliorer l'identification de l'espèce végétale.

4.6 Conclusion

Dans ce chapitre, les différentes applications de chaque étape du calcul de notre approche ont été abordées, à savoir, les applications d'extraction, de représentation et d'encodage de la macro-structure. Il était question de définir les espaces de projection de ces différentes applications.

À l'extraction de la macrostructure, on a introduit différentes représentations de l'espace d'échelle en exploitant différentes distributions des niveaux d'échelle et des points voisins. Par la suite, on a élaboré une représentation statistique de la macrostructure via la projection dans l'espace de distribution de l'information d'intensité. Cette représentation approxime la distribution en modélisant l'apparence spécifique de la surface d'écorce, et offre des propriétés d'invariance à la rotation locale tout en réduisant la sensibilité au bruit. En dernière étape, pour l'encodage de la macrostructure, deux façons différentes pour caractériser la relation entre les niveaux de l'espace d'échelle ont été présentées : un encodage radial qui encode une information graduelle sur la distribution d'information à travers les échelles de la macrostructure, et un encodage profond qui encode une information globale sur la distribution d'information dans la macrostructure par rapport à un prototype intrinsèque adaptatif, soit par résolution, soit par motif. Ces deux encodages génèrent chacun un code binaire caractérisant le macro-motif englobé par la macrostructure.

À travers ce schéma de représentation de la texture d'écorce, une famille de descripteurs statistiques des macro-motifs binaires est dérivée. Chacun de ces descripteurs représente un aspect particulier de la structure d'écorce. De nombreuses propriétés de ces descripteurs ont été présentées. D'ailleurs, mesurer les performances et l'efficacité de ces descripteurs pour l'identification des espèces végétales à partir de leurs écorces fera l'objet de la partie qui s'en suit.

Chapitre 5

Bases d'écorce et protocole d'évaluation

Sommaire

5.1 Bases d'écorce	86
5.1.1 La base Trunk12	86
5.1.2 La base AFF	86
5.1.3 La base Bark101	87
5.1.4 La base BarkNet	87
5.1.5 Autres bases d'écorce	88
5.2 Classification et mesures d'évaluation	89
5.2.1 Les classifieurs	89
5.2.2 Mesures d'évaluation	89
5.3 Méthodes en comparaison	91
5.4 Conclusion	92

AFIN d'évaluer les performances des descripteurs proposés, et valider l'efficacité de l'approche de modélisation de la texture d'écorce à base de macrostructures, il est primordial d'effectuer une série d'expérimentations sur différentes bases d'écorce. Ces expérimentations permettent de mesurer la pertinence des attributs de texture afin d'attester de leur pouvoir discriminant dans le cadre d'un système d'identification des espèces végétales à partir d'images d'écorce.

Dans ce chapitre, nous exposons le contexte d'évaluation des descripteurs statistiques de la texture d'écorce, en commençant par une présentation des bases d'écorce utilisées pour l'évaluation, ainsi que le protocole d'identification avec les différents classifieurs et les différentes mesures de performance utilisées afin d'interpréter les résultats.

5.1 Bases d'écorce

Pour l'évaluation de notre approche, **quatre** différentes bases d'écorce sont utilisées. Il s'agit des bases Trunk12 [SVAB \[2014\]](#), AFF [WENDEL et collab. \[2011\]](#), Bark101 [RATAJCZAK et collab. \[2019\]](#), et BarkNet [CARPENTIER et collab. \[2018\]](#). Le Tableau 5.1 regroupe les informations générales sur ces bases, et la Figure 5.1 présente des exemples des espèces d'écorce pour les différentes bases.

5.1.1 La base Trunk12

La base Trunk12¹ [ŠULC et MATAS \[2013\]](#) comporte 12 espèces d'écorce de la Slovénie. C'est la première base publique dédiée à l'identification des écorces.

Les images d'écorce ont été acquises sous des conditions plus ou moins contrôlées, dont on cite ce qui suit : toutes les images d'écorce ne contiennent pas de mousse, elles sont prises à une distance constante de 20cm du tronc de l'arbre avec le même appareil photo, et sous des conditions similaires de luminance. Cependant, le problème de similarité inter-espèces est bien présent pour la plupart des espèces d'écorce comme montré dans la Figure 5.1a.

La taille initiale des images de la base Trunk12 étant très grande, i.e., 3000 × 4000, toutes les images sont redimensionnées à une valeur constante de 1000 *pixels* avec préservation du ratio suivant [ŠULC et MATAS \[2013\]](#).

5.1.2 La base AFF

Parmi les travaux récents sur l'identification de l'écorce, la base la plus citée est AFF [SIXTA \[2011\]](#); [ŠULC et MATAS \[2013\]](#); [SULC et MATAS \[2014\]](#); [ŠULC et MATAS \[2017\]](#).²

AFF est une collection de 11 espèces d'écorce différentes présentes en Autriche, avec un nombre variable d'images par espèce (cf. Tableau 5.1). La Figure 5.1b montre des échantillons de chaque espèce. Les images de AFF montrent une forte similarité inter-espèces. De plus, des problèmes d'imagerie et des problèmes naturels comme la mousse, le lichen, et les blessures sont présents, faisant d'elle une base défiante pour notre application d'identification.

1. <http://www.vicos.si/Downloads/TRUNK12>

2. Cette base est fournie par 'Osterreichische Bundesforste AG' (qui donne *Austrian Federal Forestry (AFF)*).

La base AFF nous est partagée par le laboratoire 'The computer vision Lab, TU', Vienne pour des fins de recherche.

De façon similaire aux images Trunk12, nous redimensionnons les images AFF, initialement ayant une taille de $756 - 3465 \times 486 - 3264$, à une valeur constante de largeur égale à 1000pixels avec préservation du ratio suivant ŠULC et MATAS [2013]

5.1.3 La base Bark101

La base Bark101³ est une collection de 2586 images d'écorce extraites du challenge d'identification de plantes PlantNet 2017⁴ qui fait partie de la campagne d'évaluation ImageCLEF⁵. Les images d'organes végétaux de PlantNet 2017 proviennent de différentes régions de France, et elles ont été collectées à travers l'application mobile Pl@ntNet⁶. Les images d'écorce sont manuellement segmentées par RATAJCZAK et collab. [2019] afin de ne garder que l'objet écorce isolé du fond naturel, car les images initiales étaient prises dans des conditions non contrôlées. Les images d'écorce de Bark101 n'ont pas une résolution uniforme à la suite de la segmentation (des tailles variant de $69 - 800 \times 112 - 804$). Ceci produit une base avec des images qui simulent des conditions d'application réelle RATAJCZAK et collab. [2019]. Les auteurs ont également partagé l'ensemble d'entraînement et de test, où 50% des images d'une classe sont choisies pour l'entraînement.

Bark101, contient un très grand nombre d'espèces d'écorce, i.e., 101 classes différentes avec un faible nombre d'espèces par classe. Cette base présente un réel défi pour la classification des espèces d'écorce car en plus du nombre des classes, on trouve aussi une importante variabilité intra-espèce.

5.1.4 La base BarkNet

CARPENTIER et collab. [2018] a proposé une collection d'images d'écorce, nommée BarkNet 1.0⁷ à partir d'arbre de parcs et forêts de la région du Québec au Canada. BarkNet comporte 23616 images d'écorce appartenant à 23 espèces différentes. Lors de la collecte des images d'écorce, les auteurs ont pris plusieurs images du même arbre pour une espèce donnée. Le but était d'accélérer le processus de collecte d'un très grand nombre d'images, mais aussi d'incorporer une diversité d'images selon l'emplacement naturel de l'arbre. Selon ce dernier, l'apparence de l'écorce diffère et affecte l'application d'identification de son espèce végétale.

Par ailleurs, les images d'écorce sont délibérément prises sous des conditions non contrôlées de l'échelle avec des distances variables entre $20 - 60\text{cm}$ du tronc d'arbre, et à des hauteurs variables aussi. De plus, des variations de luminances aussi sont présentes telles que les images ont été prises sous différentes conditions météo sur une longue période en été par quatre caméras mobiles.

Ceci fait de la base BarkNet une base très défiante à la fois pour l'évaluation des performances des descripteurs sous plusieurs variations photométriques et naturelles, et aussi à leurs extension sur une très large collection d'espèces d'écorce.

3. <http://eidolon.univ-lyon2.fr/~remi1/Bark-101/>

4. <https://www.imageclef.org/lifeclef/2017/plant>

5. <http://www.imageclef.org>

6. <https://plantnet.org/>

7. <https://github.com/ulaval-damas/tree-bark-classification>



FIGURE 5.1 – Exemples d'écorce de chaque classe des bases d'écorce.

TABLEAU 5.1 – Information générale sur les bases d'écorce.

Base d'écorce	Trunk12	AFF	Bark101	BarkNet
#images	393	1182	2586	23616
#classes	12	11	101	23
#image/classe	30-45	16-213	2-138	64-2724
Taille d'image	1334×1000	(478-1812)×1000	(69-800)×(112-804)	224×224
Rotation	×	✓	×	×
luminance	×	✓	✓	✓
Échelle	×	✓	✓	✓
Occlusion	×	✓	✓	✓
Problèmes naturels	×	✓	✓	✓

Pour les images de la base BarkNet, on reprend les mêmes prétraitements effectués par **CARPENTIER et collab. [2018]**, i.e., toutes les images sont redimensionnées à $224 \times 224 px$. Pour le partage de la base en sous-ensembles entraînement et test, on garde le même partage des auteurs avec 80% des arbres pour l'entraînement et 20% pour le teste. Le partage de la base BarkNet est réalisé par rapport au nombre d'arbres et non pas au nombre d'images par classe, ce qui ramène le nombre de classes considérées à 20 classes au lieu de 23 [**CARPENTIER et collab., 2018**].

5.1.5 Autres bases d'écorce

De nombreux travaux sur l'identification des plantes à partir des images d'écorce ont été évalués sur des collections d'images acquises par les auteurs eux mêmes. Cependant, bien que les bases d'images comportent différentes espèces végétales, elles ne sont ni partagées ni librement accessibles. Ceci limite leur réutilisation dans le cadre d'une éventuelle évaluation d'un système

d'identification d'écorce. On cite dans ce contexte les bases de : [CHI et collab. \[2003\]](#) (8 espèces et 200 images), [WAN et collab. \[2004\]](#) (9 espèces et 160 images), [SONG et collab. \[2004\]](#) (18 espèces et 180 images), [HUANG \[2006\]](#); [HUANG et collab. \[2006b,c\]](#) (17 espèces et 300 images), [HUANG et collab. \[2006d\]](#) (24 espèces et 384 images), [LAMEI et collab. \[2010\]](#) (20 espèces et 200 images), [TSCHIEDEL GUEDES et BORGES \[2018\]](#) (51 espèces et 1836 images), [KIM et collab. \[2011\]](#) (16 espèces), et [BRESSANE et collab. \[2016, 2015b\]](#) (8 et 11 espèces avec un total d'images de 864 et 1188 respectivement).

5.2 Classification et mesures d'évaluation

5.2.1 Les classifieurs

Deux classifieurs sont utilisés sur des histogrammes des descripteurs normalisés nH :

- Le classifieur du plus proche voisin (*k-Nearest Neighbor*) k – NN ($k=1$) avec la distance L1 : 1 – NN est parmi les classifieurs les plus utilisés. En effet, il est facile à implémenter, avec un faible coût de calcul, cependant, relatif à la taille des descripteurs et leurs nombres. De plus, k – NN est non paramétrique, ce qui ne nécessite pas une étape d'apprentissage de l'ensemble des paramètres adéquats pour chaque base de test. La distance L1 entre deux descripteurs nH_x nH_y de dimension D est donnée par ce qui suit :

$$L1(nH_x, nH_y) = \sum_{i=1}^D |nH_{x_i} - nH_{y_i}| \quad (5.1)$$

De plus, on emploie le 1 – NN avec une stratégie de validation croisée, *LeaveOneOut*(LOO), où chaque image est considérée comme test et le reste des images de la base comme ensemble d'entraînement. Le descripteur de l'image requête est comparé avec tous les descripteurs des images de la base. La classe de l'image requête est déterminée par la classe de l'image réponse la plus similaire en fonction d'une mesure de similarité i.e., la distance.

- Le classifieur des machines à vecteurs supports (SVM) : SVM est un classifieur paramétrique largement utilisé dans les applications de classification et de reconnaissance. Dans notre contexte on utilise un SVM linéaire SVM_{lin} , et un SVM avec un noyau à base radiale, SVM_{rbf} comme dans [RATAJCZAK et collab. \[2019\]](#). Le noyau *rbf* est donné par :

$$K(nH_x, nH_y) = \exp(-\gamma(nH_x - nH_y)^2) \quad (5.2)$$

On utilise la bibliothèque publique LIBSVM [CHANG et LIN \[2011\]](#) avec la stratégie *One – vs – All* pour la classification multi-classes. Une étape d'optimisation des paramètres C et γ est requise pour les valeurs optimale correspondantes à un meilleur taux de classification. On applique une grille de recherche dans l'intervalle $[2^{-5} - 2^{18}]$ pour le paramètre C, et une grille dans l'intervalle $[2^{-15} - 2^8]$ pour le paramètre γ avec un pas de 2.

5.2.2 Mesures d'évaluation

Afin de quantifier l'efficacité de nos descripteurs, nous utilisons le taux de classification (accuracy) et la MAP pour étudier le comportement général de nos descripteurs sur la globalité d'une

base de test donnée. De plus, et pour une étude plus détaillée, nous proposons de visualiser le comportement de nos descripteurs sur des classes (ou bien espèces) spécifiques avec les matrices de confusion.

Score d'identification

Le score d'identification (ou bien l'accuracy de la classification) et le rapport entre le nombre des images correctement identifiées et le nombre total des images d'une base d'écorce donnée.

Pour un classifieur 1-NN-LOO avec la validation *LeaveOneOut*(LOO), le principe consiste à prendre une image de test à partir d'une base de B images, et la comparer avec le reste des images B-1. Une fonction binaire est ensuite utilisée pour déterminer si la classe de l'image retournée par 1-NN correspond à la classe original de l'image requête, on attribue 1, sinon 0. Le score final est la moyenne sur toutes les images de la base.

Pour les classifieurs 1-NN et SVM, on procède par le partage de la base en un sous ensemble d'entraînement et un autre sous ensemble de test. L'ensemble d'entraînement est utilisé pour l'apprentissage des paramètres du classifieur, comme le cas du SVM. Ensuite, la classe de chaque image de l'ensemble du test est prédite par le classifieur, et une fonction binaire détermine si la classe prédite correspond à la classe réelle de l'image de test. Le taux de classification est la moyenne des réponses binaires pour toutes les images de l'ensemble du test.

Moyenne des précisions moyennes MAP

Une autre mesure classique dans le contexte d'évaluation des systèmes de recherche par contenu est la moyenne des précisions moyennes (*Mean Average Precision MAP*). La MAP permet d'évaluer la qualité du classement des réponses retournées par le système. Elle est définie comme la valeur moyenne des précisions $APre(q)$ pour chaque requête q d'un ensemble de requêtes Qry .

$$MAP = \frac{\sum_{q=1}^{|Qry|} APre(q)}{|Qry|}; \quad (5.3)$$

La valeur de $APre(q)$ est calculée par la formule suivante :

$$APre(q) = \frac{\sum_{k=1}^n (Pre(k)rel(k))}{\#d'images\ pertinente\ pour\ q} \quad (5.4)$$

où :

q est le nombre des images retournées.

$Pre(k)$ est la précision à la valeur k dans la liste des images retournée, i.e., $Pre(k)$ est la proportion des images pertinentes retournées parmi les k images retournées.

$rel(k)$ est une fonction binaire qui mesure la pertinence de l'image retournée au rang k .

La MAP permet de mesurer la précision du système de recherche et de l'évaluer par une seule valeur. Elle prends des valeurs dans l'intervalle $[0, 100]$, où, plus les valeurs tendent vers 100, plus le système est performant.

Matrice de confusion

La matrice de confusion permet de visualiser de plus près si le système d'identification arrive à identifier correctement les différentes classes en les séparant les unes des autres. Les lignes de la matrice de confusion représentent les nombres réels des images d'écorce par classe, et les colonnes représentent les nombres estimés par le classifieur.

La matrice de confusion est un outil simple et efficace pour déterminer les classes que le système d'identification n'arrive pas à différencier. Ceci permet d'apporter des explications et des interprétations au comportement de notre approche.

D'autres mesures peuvent être tirées de la matrice de confusion comme les valeurs de précision et rappel globales ($Rappel_g$, $Précision_g$), et la mesure F ($F1 - score$) comme suit :

$$\begin{aligned} Rappel_g &= \frac{tp}{tp + fp} \\ Précision_g &= \frac{tp}{tp + fn} \\ F1 - score &= \frac{2tp}{2tp + fp + fn} \end{aligned} \quad (5.5)$$

où : tp : vrai positif; fp : faux positif; fn : faux négatif. Les valeurs de la $F1 - score$ sont définies entre $[0, 1]$, plus c'est proche de 1 mieux c'est.

5.3 Méthodes en comparaison

Pour évaluer l'efficacité et la pertinence de notre approche et la capacité des descripteurs statistiques à caractériser la texture des espèces d'écorce, nous avons implémenté, ou utilisé l'implémentation partagée par les auteurs, des différentes méthodes basées sur un encodage de motifs binaires. Les méthodes en comparaison sont présentées dans le Chapitre 2.

En terme de catégorie des méthodes utilisées pour la comparaison, c'est des méthodes structurales basées sur un encodage LBP, faisant de cette catégorie de méthodes la plus compétitive avec l'approche proposée dans ce travail. D'autres méthodes de description de texture sont également prises en considération lors de l'évaluation comparative de nos descripteurs. Il s'agit là d'une comparaison avec différentes représentations de la texture, comme les matrices de cooccurrences (approche statistique) et les filtres de Gabor (approche multi-échelles et multi-résolutions).

Afin de rendre la comparaison équitable et significative, toutes les méthodes en comparaisons sont en multi-échelles selon la représentation de l'espace d'échelle établie par leurs auteurs originaux, dans le cas échéant, le schéma multi-échelles de MResLBP [OJALA et collab. \[2002\]](#) est appliqué. De plus, le mapping *riu2* est utilisé pour toutes les approches à base de motifs binaires. En effet, l'utilisation de ce mapping est justifiée par deux principales raisons : (i) C'est le mapping le plus utilisé dans la plupart des méthodes en comparaison. (ii) ce mapping rend la comparaison plus équitable en terme d'invariance aux changements de rotation, et de la compacité du descripteur. Ceci est tout à fait compatible avec nos descripteurs qui ont la propriété d'être invariants aux changements de rotation locale et d'une dimension réduite et compacte.

En dernier lieu, lorsqu'il n'y a pas référence à l'utilisation d'un espace de couleur spécifique, les descripteurs des méthodes en comparaisons sont extraits à partir des images en niveaux de

gris.

5.4 Conclusion

Ayant exposé dans ce chapitre l'approche proposée pour l'évaluation de nos descripteurs, qui porte essentiellement sur l'ensemble des bases d'écorce, le protocole de classification, et les méthodes en comparaison, on présente dans la suite des chapitres les résultats expérimentaux sur les différentes bases d'écorce.

Chapitre 6

Résultats expérimentaux

Sommaire

6.1 Étude comparative des variantes LBP multi-échelles	94
6.1.1 Contexte de l'étude	94
6.1.2 Résultats comparatifs	95
6.1.3 Discussion	99
6.2 Évaluation des descripteurs de macro-motifs binaires statistiques	100
6.2.1 Résultats sur la base Trunk12	101
6.2.2 Résultats sur la base AFF	112
6.2.3 Résultats sur la base Bark101	120
6.2.4 Résultats sur la base BarkNet	126
6.3 Classification des écorces par les CNN	133
6.3.1 Détails d'implémentation	133
6.3.2 Classification d'écorce avec ResNet34	134
6.3.3 Comparaison entre SRBP, SMBP et ResNet34	135
6.4 Conclusion	138

DANS ce chapitre, nous présentons les résultats expérimentaux des descripteurs proposés dans le cadre de l'identification automatique des espèces végétales à partir des images d'écorce. Pour cela, l'étude expérimentale est principalement divisée en trois parties. La première partie porte sur une comparaison quantitative des variantes multi-échelles à base de motifs binaires LBP pour la description de la texture d'écorce. L'objectif de cette étude comparative est d'étudier l'applicabilité des motifs multi-échelles LBP pour la caractérisation de la structure d'écorce. En effet, cette comparaison sert de référence pour valider notre choix de l'encodage LBP pour la description des structures d'écorce, notamment avec un schéma multi-résolutions. La seconde partie porte essentiellement sur l'évaluation de nos descripteurs statistiques des macro-motifs binaires pour la description de la texture d'écorce à savoir, les descripteurs radiaux $SRBP_m$ et les descripteurs profonds $SMBP_m$. Pour cette partie, on procède par une évaluation des paramètres optimaux de chaque descripteur, après on présente les résultats comparatifs avec les méthodes de l'état de l'art sur chacune des bases d'écorce présentées dans le Chapitre 5. Au final, la dernière partie porte sur la classification des espèces d'écorce par un modèle CNN, le ResNet34 [HE et collab. \[2016\]](#). Le choix du modèle CNN est essentiellement basé sur l'étude de [CARPENTIER et collab. \[2018\]](#). Dans cette partie, il s'agit aussi de présenter un comparatif entre nos méthodes conçues selon le schéma classique (*handcrafted*) avec une méthode à base d'apprentissage profond.

6.1 Étude comparative des variantes LBP multi-échelles

Dans ces expériences, nous visons à évaluer différentes variantes LBP multi-échelles pour la représentation de la structure d'écorce à des niveaux d'échelle élevés. L'étude comparative porte essentiellement sur la validation d'une représentation multi-échelles pour la description de la texture de l'écorce.

6.1.1 Contexte de l'étude

Dans le cadre de l'évaluation de la description multi-échelles de la texture d'écorce à base de motifs LBP, nous avons ré-implémenté ou utilisé l'implémentation existante de différentes variantes pouvant générer des descriptions comparables de la surface de l'écorce. La plupart de ces méthodes ont été testées sur des images de textures génériques relevant de la surface d'objets naturels mais aussi d'objets synthétiques. Afin de mettre en œuvre un contexte de comparaison significatif, nous avons utilisé une taille uniforme de l'espace d'échelle pour toutes les variantes de façon à avoir le même support spatial capturant une information de macrostructure.

Les variantes LBP multi-échelles retenues ne nécessitent pas un apprentissage de l'ensemble de paramètres optimaux afin de produire une description spécifique pour chaque espèce d'écorce. Par conséquent, leur applicabilité sur des images d'écorce est facilement généralisée. La plupart des variantes en comparaison sont basées sur l'utilisation des filtres de lissages (filtres passe-bas) et suivent différentes conceptions de l'espace d'échelle. Cependant, le principe de représentation multi-échelles demeure classique, i.e., un descripteur LBP est calculé à chaque niveau d'échelle indépendamment des autres niveaux. Le descripteur final est généré par une concaténation de tous les descripteurs locaux LBP par échelle. Ces méthodes sont au nombre de quatre, à savoir :

TABLEAU 6.1 – Les valeurs des paramètres des différentes variantes LBP multi-échelles en fonction du niveau de l'échelle.

Méthode	LBPF		MSLBP		PLBP	MBLBP
	MÄENPÄÄ et PIETIKÄINEN [2003]		ŠULC et MATAS [2013]		QIAN et collab. [2011]	LIAO et collab. [2007]
\mathcal{L}	σ	R_i	σ	R_i	Taille image	Taille filtre
1	0.38	2.43	0.28	1.41	600 × 400	3 × 3
2	0.85	5.44	0.39	2	300 × 200	9 × 9
3	1.91	12.19	0.55	2.83	150 × 100	15 × 15
4	4.27	27.30	0.78	4	75 × 50	21 × 21
5	9.56	61.16	1.10	5.66	38 × 25	27 × 27
6	21.42	136.98	1.56	8	19 × 13	33 × 33

TABLEAU 6.2 – Valeurs de MAP pour les variantes LBP basique et MResLBP sur les bases Trunk12 et AFF.

Méthode	LBP OJALA et collab. [2002]	LBP _{u2} OJALA et collab. [2002]	MResLBP OJALA et collab. [2002]
Trunk12	31.42	29.69	30.03
AFF	31.46	31.00	32.29

LBP Filtering (LBPF) MÄENPÄÄ et PIETIKÄINEN [2003], Multi-Scale LBP (MSLBP) ŠULC et MATAS [2013], Pyramid-based LBP (PLBP) QIAN et collab. [2011], et Multi-Block LBP (MBLBP) LIAO et collab. [2007]. Le LBP OJALA et collab. [2002] avec un voisinage local est utilisé comme référence de comparaison afin de mesurer l'impact de la représentation multi-échelles sur les images d'écorce. Le MResLBP OJALA et collab. [2002] sert de référence pour analyser la conception de l'espace d'échelle en conjonction avec les réponses des filtres de lissage, et leurs effets sur la réduction du bruit. Ces variantes ont été présentées en détail dans le Chapitre 2.

Le Tableau 6.1 présente les valeurs des paramètres de chacune des variantes LBP multi-échelles en fonction du niveau d'échelle \mathcal{L} qui prend des valeurs dans l'intervalle $[1, 6]$.

6.1.2 Résultats comparatifs

L'identification des espèces d'écorce est réalisée par le classifieur k -NN/LOO ($k = 1$). La description de la texture d'écorce avec les variantes LBP est réalisée sur les images des bases Trunk12 et AFF représentées en niveau de gris. Pour cette étude comparative et différemment de la configuration des images d'écorce établie dans le Chapitre 5, les images des deux bases d'écorce sont redimensionnées à une résolution de 600×400 px à partir de leurs tailles initiales. Nous avons également évalué différents mappings des motifs LBP, dénotés *org* et *u2* pour le cas où tous les motifs possibles sont retenus, i.e., tous les 2^{N_i} motifs, et le cas des motifs uniformes selon la mesure d'uniformité dans OJALA et collab. [2002]. En plus, une représentation de n motifs uniformes, dénotée n , est testée pour la variante MBLBP, tel que $n = 63$ LIAO et collab. [2007].

Communément, le paramètre relatif au nombre de points voisins par échelle pour toutes les variantes est donné par $N_i = 8$, une valeur fixe et constante pour tous les niveaux d'échelle, produisant ainsi, dans le cas des motifs *org*, un descripteur final d'une dimension de $2^{N_i} \times \mathcal{L}$.

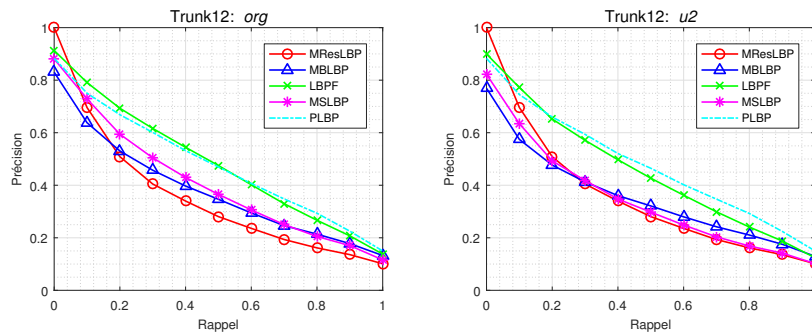
Les performances sont exprimées en terme de MAP (*Mean Average Precision*). On compare en premier lieu une description de l'écorce par le descripteur local LBP OJALA et collab. [2002] et sa description en multi-échelles MResLBP OJALA et collab. [2002]. Les valeurs de MAP pour cette première analyse sont présentées dans le Tableau 6.2.

TABLEAU 6.3 – Valeurs de MAP pour les variantes LBP multi-échelles sur la base Trunk12.

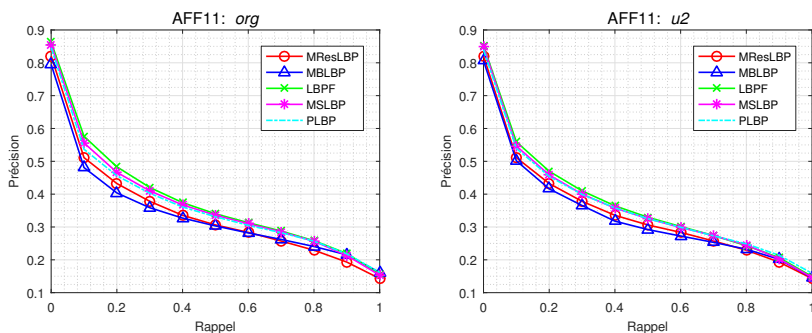
Trunk12	MBLBP			LBPF		MSLBP		PLBP	
	LIAO et collab. [2007]			MÄENPÄÄ et PIETIKÄINEN [2003]		ŠULC et MATAS [2013]		QIAN et collab. [2011]	
\mathcal{L}	org	u2	n	org	u2	org	u2	org	u2
1	32.34	28.32	24.76	32.85	31.11	31.30	28.75	36.23	34.74
2	34.43	29.16	28.75	37.44	35.51	31.62	29.43	40.93	39.67
3	35.80	29.84	31.22	42.80	40.14	32.68	30.59	45.05	44.29
4	36.46	30.74	32.58	47.66	44.28	35.09	33.15	47.10	46.74
5	36.66	31.22	33.32	39.07	36.64	37.38	35.01	37.60	38.27
6	36.57	31.79	33.77	17.38	18.20	39.59	36.66	23.45	24.85

TABLEAU 6.4 – Valeurs de MAP pour les variantes LBP multi-échelles sur la base AFF.

AFF	MBLBP			LBPF		MSLBP		PLBP	
	LIAO et collab. [2007]			MÄENPÄÄ et PIETIKÄINEN [2003]		ŠULC et MATAS [2013]		QIAN et collab. [2011]	
\mathcal{L}	org	u2	n	org	u2	org	u2	org	u2
1	32.12	30.28	26.39	34.10	33.95	32.14	31.66	34.99	33.69
2	33.86	31.75	29.40	35.62	34.96	33.25	33.08	35.17	34.18
3	33.40	31.90	29.98	36.60	35.39	34.09	33.74	34.68	34.78
4	32.74	31.72	30.03	35.50	34.19	34.87	34.53	32.93	32.78
5	32.17	31.65	30.00	29.75	28.44	35.33	34.71	27.82	27.91
6	31.67	31.53	29.93	18.97	19.91	35.80	34.77	22.71	22.98



(a) Trunk12.



(b) AFF.

FIGURE 6.1 – Courbes de Rappel/Précision des variantes LBP multi-échelles sur les bases Trunk12 et AFF.

À partir du Tableau 6.2, on peut voir que la combinaison des descripteurs LBP^{u2} sur trois niveaux d'échelle, i.e., MResLBP, contribue à améliorer la description de la texture d'écorce par rapport à une représentation locale avec un seul niveau de résolution spatiale. Cependant, il est clair que les motifs uniformes ne sont toujours pas satisfaisants pour caractériser la variation de l'information à des niveaux d'échelle élevés. Ceci est principalement dû au fait que le mapping *u2* regroupe une grande partie de l'information, considérée bruitée, comme motifs non uniformes. Toutefois, élaborer une description multi-échelles par les motifs *org*, en retenant un nombre graduellement élevé de points voisins par niveau d'échelle, produit un descripteur final de dimension conséquente ($Dim = 2^8 + 2^{16} + 2^{24}$).

Une représentation multi-échelles est plus intéressante qu'une représentation très locale. Cependant, la performance de celle là est nettement déterminée par la qualité de l'information encodée, et l'effet du bruit sur les valeurs d'intensité des points voisins distants, d'où la nécessité d'estomper les valeurs aberrantes par un lissage local via l'application de filtre passe-bas MÄENPÄÄ et PIETIKÄINEN [2003].

La comparaison des variantes multi-échelles LBP basées sur le calcul du code LBP à partir d'une image de réponse de filtre est donnée dans les Tableaux 6.4 et 6.3 pour les bases Trunk12 et AFF utilisant différents mappings des motifs LBP.

Afin de mieux visualiser et comparer les performances de ces variantes entre elles sur les différentes bases d'écorce, les courbes de Rappel/Précision sont tracées pour chacune des variantes utilisant le niveau d'échelle où la plus grande valeur de MAP est présente, (c.f les Figures 6.1a et 6.1b).

Globalement, les valeurs de la MAP obtenues par les différentes méthodes LBP sur la base Trunk12 sont relativement supérieures à celles obtenues pour la base AFF. Ceci est justifié par la taille de la base, l'homogénéité des classes d'espèce, ainsi que la présence d'un certain nombre de problèmes d'imagerie, et problèmes naturels pouvant altérer la qualité de l'image. En effet, la base AFF recense une variation considérable d'apparence entre des images d'écorce appartenant à la même espèce, tandis que les classes d'espèce de la base Trunk12 sont assez homogènes.

Le bénéfice de l'application des filtres passe-bas avant le calcul du code LBP est significativement notable par rapport à la représentation sans filtre de MResLBP OJALA et collab. [2002]. Représenter la valeur d'intensité d'un point voisin par une moyenne pondérée de son voisinage réduit considérablement l'effet du bruit.

Les performances rapportées confirment que les valeurs de MAP ont tendance à augmenter en fonction de la taille de l'espace d'échelle, en l'occurrence sur la base Trunk12, les valeurs de MAP de LBPF (*org*) passe de 32.85 avec $\mathcal{L} = 1$ à 47.66 avec $\mathcal{L} = 4$, de même sur la base AFF, où on trouve une MAP de 34.1 à $\mathcal{L} = 1$ et une MAP de 36.6 à $\mathcal{L} = 3$ avec le même descripteur. De plus, les résultats avec les motifs *org* sont assez meilleurs que ceux obtenus avec les motifs uniformes *u2*. En effet, pour les motifs uniformes, le même poids est attribué à une partie de l'information considérée non uniformes, ce qui peut cumuler une perte d'information de texture. La représentation d'uniformité selon LIAO et collab. [2007] avec $n = 63$ repose sur la distribution statistique des motifs LBP dans l'image (leurs fréquences) et non pas sur des propriétés de texture (comme pour la mesure U de OJALA et collab. [2002]). Ceci peut mener à une confusion de mise en correspondance des histogrammes MBLBP d'une paire d'images requête-réponse justifiant ainsi les

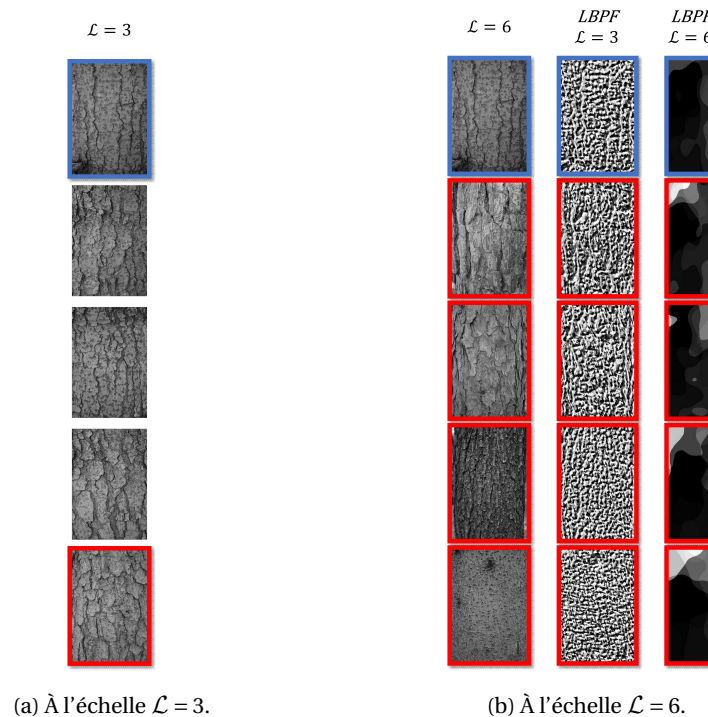


FIGURE 6.2 – Visualisation des images LBP des réponses retournées à une image d'écorce requête.

faibles valeurs de MAP pour la base AFF. Quand à la représentation des images de la base Trunk12, les n valeurs fréquentes et uniformes de MBLBP ont coïncidé avec l'espace de projection lors de la mise en correspondance entre les descripteurs des paires d'images requête-réponse.

En ce qui concerne le type du filtre passe-bas, le filtre Gaussien réduit considérablement l'effet de bruit, i.e., LBP [MÄENPÄÄ et PIETIKÄINEN \[2003\]](#), MSLBP [ŠULC et MATAS \[2013\]](#), et PLBP [QIAN et collab. \[2011\]](#), permettant d'obtenir des scores MAP clairement meilleurs que ceux obtenus avec le filtre moyen, i.e., MBLBP [LIAO et collab. \[2007\]](#). Cependant, il est intéressant de faire une analyse détaillée de la représentation de l'espace d'échelle pour ces trois variantes.

L'espace d'échelle de MSLBP [ŠULC et MATAS \[2013\]](#) est la représentation fine de celui de LBP [MÄENPÄÄ et PIETIKÄINEN \[2003\]](#). Par conséquent, les valeurs de MAP de MSLBP obtenues à l'échelle $\mathcal{L} = 6$ avec un rayon maximal de $R_6 = 8$ sont approximatives à celles obtenues par LBP à l'échelle $\mathcal{L} = 3$ et un rayon maximal de $R_3 = 12.19$ (cf. Tableau 6.1). Ceci laisse deviner que les performances de MSLBP vont augmenter d'avantage pour des espaces d'échelle plus grands. Cependant, une telle représentation affecte directement la dimension finale du descripteur, tenant compte de la combinaison de multiples descripteurs multi-échelles [ŠULC et MATAS \[2013\]](#).

La limite de la représentation de l'espace d'échelle par une distribution logarithmique des niveaux d'échelle, i.e., LBP [MÄENPÄÄ et PIETIKÄINEN \[2003\]](#), est marquée par une rapide dégradation des valeurs de MAP à partir de l'échelle $\mathcal{L} = 3$. Afin de visualiser ce comportement de plus près, nous avons examiné la représentation des images d'écorce par un descripteur LBP à l'échelle $\mathcal{L} = 3$, et à l'échelle $\mathcal{L} = 6$. La Figure 6.2 montre le résultat de recherche des images similaires à une image d'écorce de la base AFF, et leurs représentations correspondantes par des codes LBP.

Il est visiblement clair que plus l'espace d'échelle devient grand, plus la valeur σ propre au filtre

Gaussien augmente (cf. Tableau 6.1). Ceci veut dire que lors du calcul des moyennes pondérées des régions effectives, toutes les valeurs dépassant le seuil σ sont floutées (ou bien estompées), causant une perte considérable d'information. Par conséquent, ceci introduit une forte confusion inter-espèces, comme le montre les représentations LBPF à l'échelle $\mathcal{L} = 6$ des différentes espèces d'écorce considérées comme similaires à l'espèce requête (cas d'un faux positif).

En ce qui revient pour une représentation de l'espace d'échelle par une décomposition en pyramide Gaussienne, i.e., cas de PLBB QIAN et collab. [2011], on remarque un comportement similaire à celui de LBPF. En effet, en plus de la représentation très floue de la structure de texture engendrée par le filtre Gaussien à des niveaux d'échelle élevés, le sous-échantillonnage de l'image en entrée affecte la quantité d'information à encoder à partir des résolutions très faibles.

6.1.3 Discussion

À travers cette étude comparative des variantes LBP multi-échelles sur deux bases d'écorce différentes, les points suivant sont retenus :

Une description multi-échelles est plus adaptée pour la description des images de texture d'écorce qu'une description locale LBP. En effet, ceci coïncide avec la description spécifique de la surface d'écorce en terme de macrostructure, et est confirmé par les résultats expérimentaux obtenus lors de la comparaison entre LBP et MResLBP. Une représentation multi-échelles permet d'encoder une information sur la macrostructure avec des supports spatiaux plus larges. Cependant, il convient d'utiliser l'ensemble des motifs possibles (notés *org*) pour recouvrir l'information négligée par les motifs uniformes. Toutefois, il est important d'appliquer un mécanisme de réduction de bruit lorsque l'espace d'échelle s'étale sur de larges supports spatiaux.

On a trouvé qu'une représentation du voisinage par des moyennes pondérées (moyenne de bloc ou moyenne Gaussienne) enregistre des performances plus élevées que celles retournées par MResLBP où les valeurs d'intensité brutes des points voisins sont utilisées directement. En effet, d'un côté, un certain nombre de points voisins est nécessaire pour représenter le voisinage multi-échelles, de l'autre, ce nombre est relatif à la taille du niveau d'échelle (le rayon) ce qui conduit à une représentation excessive du voisinage par des valeurs d'intensité plus ou moins bruitées. Les approches basées sur une représentation de voisinage par moyenne pondérée limitent pour la plupart le nombre de points voisins par échelle, et attribuent des fenêtres de filtre de taille relative au rayon pour collecter l'information autours des pixels distants. Par conséquent, le voisinage multi-échelles est représenté par des valeurs d'intensité beaucoup moins sensibles au bruit.

Parmi les approches multi-échelles à base de filtres, les approches à base de filtres Gaussiens, i.e., LBPF, MSLBP, et PLBP, sont plus efficaces que l'approche MBLBP à base de filtre moyen. En effet, le calcul de la moyenne locale est affecté par les valeurs extrêmes dans le voisinage. Dans ce cas, la conception de l'espace d'échelle par des filtres Gaussiens devient plus intéressante. Cependant, le filtre Gaussien couplé avec une décomposition en pyramide pour une analyse multi-résolutions, i.e., PLBP, atteint vite sa limite lorsque la taille de l'image est très petite mais aussi très floutée à des niveaux élevés de la pyramide.

Une conception grossière de l'espace d'échelle, LBPF, est plus performante qu'une conception fine, MSLBP. En effet, ceci permet de représenter une information distante qui s'avère plus pertinente pour la description des structures d'écorce.

Néanmoins, le calcul des images réponses de filtres à des niveaux d'échelle élevés contribue à introduire une confusion entre les descriptions de différentes espèces d'écorce, où une similarité inter-espèces due à la perte conséquente de l'information par des filtres Gaussiens à des niveaux élevés est remarquablement notée. De ce fait, il est nécessaire de trouver un compromis entre atténuer l'effet du bruit à des niveaux d'échelle élevés et représenter une information distante pertinente.

Cette étude comparative nous a permis d'un côté de valider la performance d'une représentation multi-échelles LBP pour la description de la structure d'écorce, et de l'autre côté, de la nécessité d'appliquer un schéma d'encodage qui tient compte à la fois de l'effet du bruit et la taille du descripteur lors de l'extension sur des supports spatiaux de tailles beaucoup plus importantes.

Ces contraintes sont pour la plupart prises en considération dans l'approche proposée dont la validation de sa performance sera examinée dans la partie qui suit.

6.2 Évaluation des descripteurs de macro-motifs binaires statistiques

Dans cette seconde partie de l'étude expérimentale, on évalue les descripteurs statistiques binaires sur les différentes bases d'écorce. L'étude expérimentale est principalement composée de deux parties : une partie évaluative et une partie comparative comme suit :

1. Étude évaluative : Cette partie sert à valider la performance d'une description de texture d'écorce à base de descripteurs statistiques macro-motifs en passant par les études suivantes des différents paramètres :

Prototype statistique : pour mettre en évidence le calcul d'un code macro-motif par un gradient en profondeur, et aussi valider le choix du prototype statistique :

- ▶ μ^{scale} ,
- ▶ $\mu^{pattern}$.

À priori, cette évaluation est réalisée à partir des macrostructures circulaires. Le prototype optimal retenu est par la suite utilisé pour l'évaluation des descripteurs statistiques profonds à partir des macrostructures en anneaux.

Conception de l'espace d'échelle : pour déterminer la taille de la macrostructure optimale qui permet de représenter des structures de texture caractéristiques des différentes espèces d'écorce, ceci inclus les paramètres suivants :

- ▶ \mathcal{L} : La taille de l'espace d'échelle.
- ▶ R_1 : La localité de la macrostructure.
- ▶ τ : La distance radiale entre les niveaux d'échelle, et la largeur de l'anneau.
- ▶ $AllN, OneN$: La distribution des points voisins.

Combinaisons des descripteurs statistiques : afin d'étudier l'influence des différents descripteurs statistiques, seuls, et leurs différentes combinaisons.

2. Étude comparative : Cette partie présente une comparaison de nos descripteurs avec des approches de l'état de l'art dans le contexte de l'identification des espèces végétales à partir d'images d'écorce.

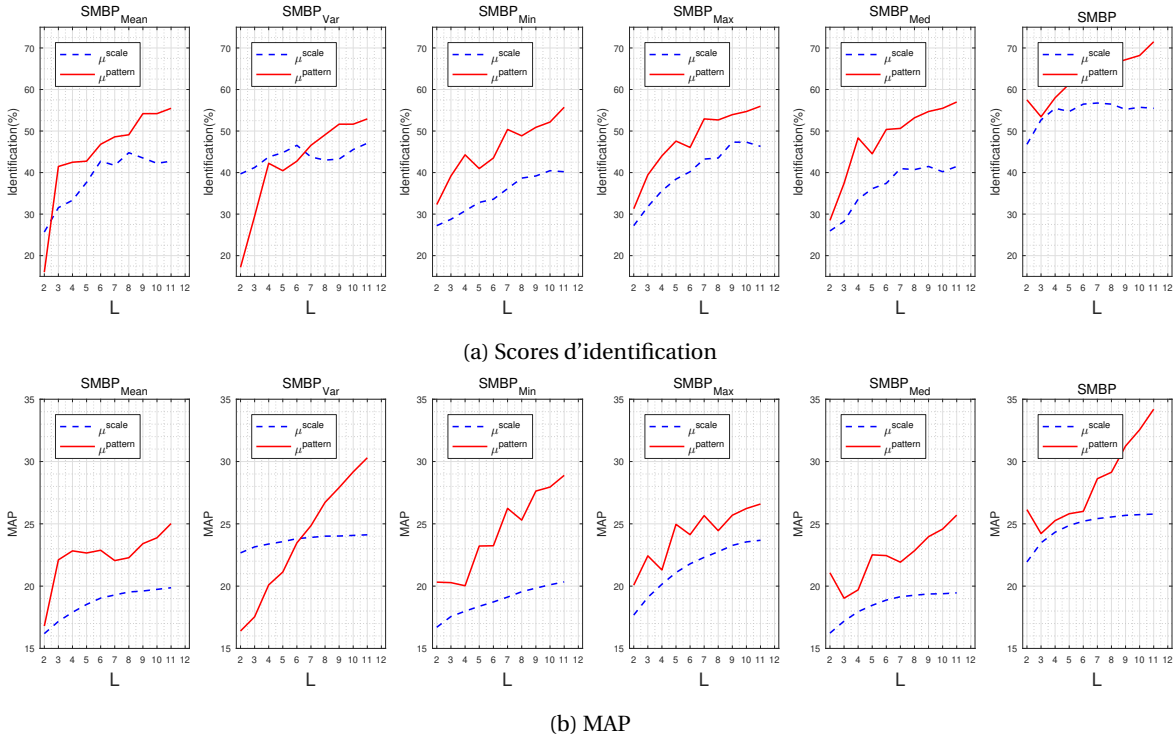


FIGURE 6.3 – Taux d'identification (%) et valeurs MAP des descripteurs singuliers $SMBP_m$ et leur combinaison $SMBP$ avec les prototypes statistiques μ^{scale} et $\mu^{pattern}$ sur la base Trunk12. La configuration utilisée est $\{R_1 = 1, \tau = 1, AllN\}$

6.2.1 Résultats sur la base Trunk12

6.2.1.1 Évaluation des paramètres

6.2.1.1.1 Prototypage statistique

La Figure 6.3 présente les taux d'identification et les valeurs MAP pour les descripteurs $SMBP_m | m \in \{\mathcal{M}\}$ et leur combinaison, $SMBP$, sur la base Trunk12.

Ces descripteurs sont calculés avec le prototype par échelle μ^{scale} , et le prototype par motif $\mu^{pattern}$. La configuration utilisée est $\{R_1 = 1, \tau = 1, AllN\}$, $\mathcal{L} = \llbracket 2, 11 \rrbracket$,

À partir de la Figure 6.3, on peut tirer les observations suivantes :

- Les scores d'identification des $SMBP_m$ avec $\mu^{pattern}$ ou bien avec μ^{scale} augmentent graduellement en fonction de la taille de la macrostructure, donnant des valeurs élevées à $\mathcal{L} = 11$.
- Les descripteurs $SMBP_m$ basés sur le prototype local $\mu^{pattern}$ dépassent significativement leurs correspondants basés sur le prototype global μ^{scale} . L'écart devient de plus en plus grand pour des valeurs d'échelle élevées. On note un écart moyen de +11.85% entre les descripteurs $SMBP_m$ avec $\mu^{pattern}$ et ceux avec μ^{scale} à $\mathcal{L} = 11$.
- La combinaison des cinq descripteurs statistiques permet d'obtenir des scores d'identification nettement meilleurs que les descripteurs individuels quelque soit le type du prototype utilisé. Ici, l'écart entre la combinaison basée sur $\mu^{pattern}$ et celle basée sur μ^{scale} est de +16.03% à la plus grande échelle. Ceci montre l'efficacité d'un encodage à base du proto-

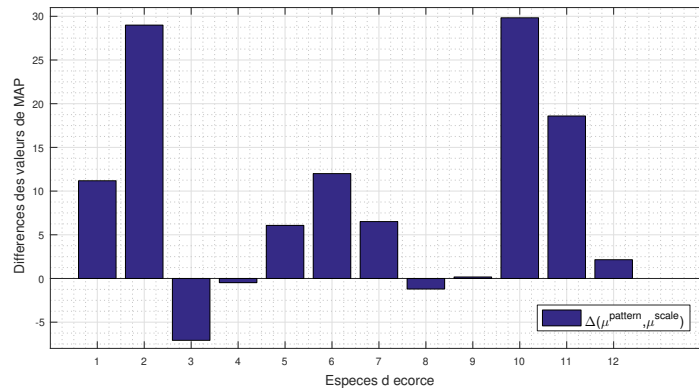


FIGURE 6.4 – Différence entre les valeurs des MAP par espèce pour le descripteur SMBP avec μ^{scale} et celle du descripteur SMBP avec $\mu^{pattern}$.



FIGURE 6.5 – Variabilité intra-espèce pour l'espèce #3 de la base Trunk12.

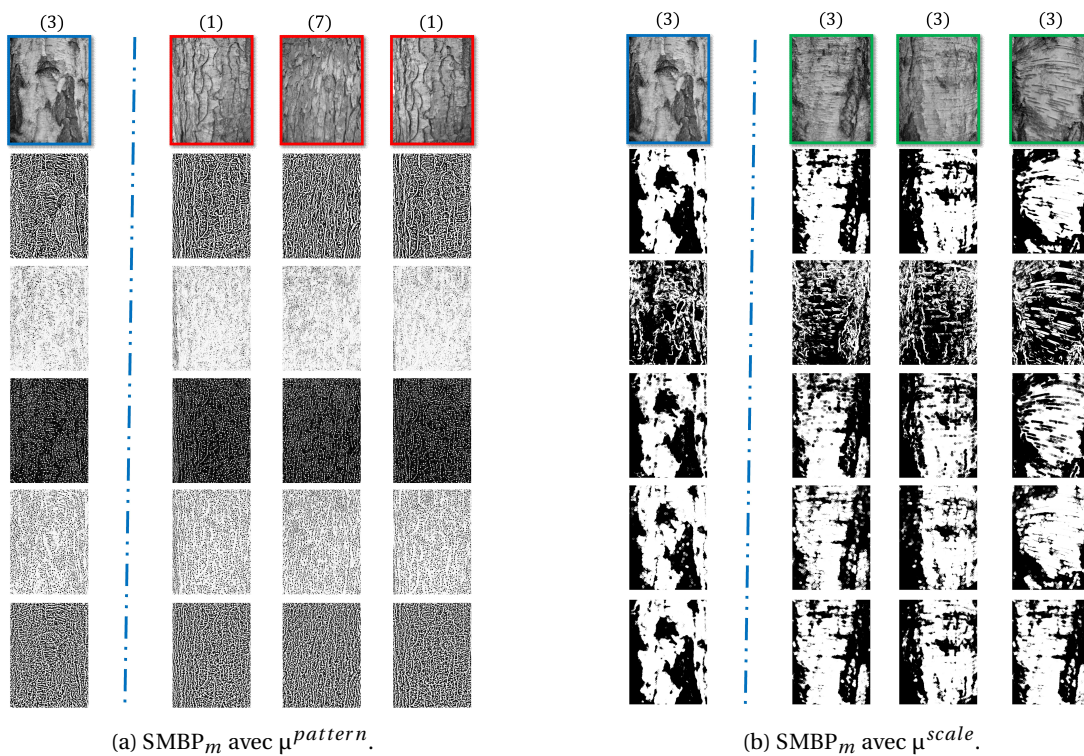


FIGURE 6.6 – Image requête de l'espace #3 et les trois premières images retournées avec leurs représentations statistiques pour les différents prototypes.

type local pour la représentation de la structure de texture d'écorce.

- Les précédentes remarques sont soutenues par les valeurs de MAP de la Figure 6.3b, dont l'évolution est représentée par les courbes des descripteurs singuliers et combinés.

Afin de voir de plus près la représentation des descripteurs statistiques des différentes macrostructures, nous examinons les taux d'identification par espèce en utilisant le descripteur global SMBP une fois avec le prototype $\mu^{pattern}$ et une fois avec le prototype μ^{scale} avec la configuration $\{\mathcal{L} = 11, R_1 = 1, \tau = 1, AllN\}$.

On désigne par la valeur $\Delta(\mu^{pattern}, \mu^{scale})$ la différence entre les valeurs des MAP par espèce pour le descripteur SMBP avec $\mu^{pattern}$ et celles du descripteur SMBP avec μ^{scale} . Ainsi, les valeurs positives de $\Delta(\mu^{pattern}, \mu^{scale})$ indiquent que le SMBP ne reconnaît pas beaucoup plus d'images avec le prototype μ^{scale} que avec le prototype $\mu^{pattern}$, et l'inverse indique que le prototype local ne permet pas de différencier les images des espèces concernées des autres de leur base.

La Figure 6.4 montre les résultats de $\Delta(\mu^{pattern}, \mu^{scale})$. De façon générale, lorsque le prototype local par motif $\mu^{pattern}$ est utilisé pour l'encodage du descripteur SMBP, presque toutes les espèces sont bien identifiées et mieux distinguées les unes des autres que lorsque le prototype global par échelle μ^{scale} est utilisé. Ceci est vérifié pour 9 espèces sur 12 avec le prototype local, tandis que le prototype global dépasse légèrement le prototype local en terme de MAP pour les images de l'espèce #3.

Pour l'espèce #3, on trouve que les images 'fausses positives' retournées par $\mu^{pattern}$ proviennent du fait que les images de cette classe montrent une forte variabilité intra-espèce (cf Figure 6.5), et que la présence d'un motif dominant n'est pas toujours vérifiée pour toutes les images de cette espèce.

La Figure 6.6 montre un exemple d'image requête de l'espèce #3 et les trois premières images retournées avec leurs images de codes statistiques avec le prototype local (cf. Figure 6.6a), et aussi avec le prototype global (cf. Figure 6.6b).

On peut voir qu'avec le prototype local, les trois premières images retournées ne sont pas de la même classe que l'image requête. D'après les images des codes statistiques, on constate que le prototype local $\mu^{pattern}$ considère que les codes macro-motifs des images de l'espèce #1 et #7 sont très similaires à ceux de l'espèce #3, et donc ont une distribution similaire de l'intensité dans le voisinage macro. Cependant, avec le prototype global, qui représente la moyenne de la distribution par rapport à une échelle donnée pour tous les pixels de l'image, les trois images retournées sont de la même classe que l'image requête, et leurs représentations statistiques sont assez similaires. Ceci correspond au fait que pour cette classe d'écorce, la distribution des macrostructures est très variable, et aussi de l'absence d'un macro-motif uniforme et caractérisant de ses images. Par ailleurs, l'espèce #3 est la seule classe non homogène de la base Trunk12 ce qui représente un cas particulier pour le comportement de $\mu^{pattern}$ qu'on ne peut pour autant généraliser.

Au final, il est important de noter que le score global sur toutes les images de la base Trunk12 avec SMBP et le prototype local $\mu^{pattern}$ reste largement plus élevé que celui obtenu par SMBP avec μ^{scale} , i.e., (71.50% contre 55.47%). De même pour les valeurs de MAP avec 34.20 pour $\mu^{pattern}$ contre 25.78 pour μ^{scale} .

Ainsi, on se basant sur les résultats obtenus dans cette étude, nous choisissons d'utiliser le prototype intrinsèque local $\mu^{pattern}$ lors du calcul du code macro-motif pour les descripteurs SMBP_m dans la suite des expérimentations, pour toutes les bases d'écorce, aussi bien pour les macrostructures circulaires que celles en anneaux.

TABLEAU 6.5 – Taux d'identification (%) du descripteur circ-SRBP sur la base Trunk12 avec différentes représentations de la macrostructure.

circ-SRBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	57.25	62.08	61.06	61.32	64.37	65.13	65.13	65.90	66.66	66.92	60.05	60.81	63.35	63.10	64.37	65.39	67.43	68.70	59.54	64.63	64.37	64.63	65.13	67.68
$R_1 = 2$	54.45	57.76	60.30	62.84	63.10	64.88	65.90	66.92	67.68	70.22	54.19	60.30	62.08	62.84	65.90	67.43	68.70	70.22	56.74	60.81	62.59	66.92	69.21	70.73
$R_1 = 3$	52.92	54.96	56.23	60.05	63.86	65.13	65.13	66.41	67.68	69.21	52.92	56.99	60.30	63.10	64.88	66.92	68.95	70.22	53.94	59.03	62.84	63.35	66.66	67.43
$R_1 = 4$	52.67	54.70	57.25	63.86	65.13	65.13	65.64	67.17	67.43	68.70	51.65	57.50	60.81	63.61	67.17	67.93	69.46	69.97	52.92	60.81	64.63	67.17	68.44	71.24
$R_1 = 5$	52.41	55.97	58.01	60.30	63.61	65.13	66.41	67.17	68.95	69.46	52.67	56.99	61.57	64.37	66.92	65.90	69.21	70.73	55.47	62.59	65.13	67.68	68.95	71.24
$R_1 = 6$	58.01	61.06	63.10	63.10	65.64	67.68	68.44	68.19	69.21	69.21	56.99	60.30	63.86	64.37	66.92	69.21	69.72	70.48	58.26	61.83	66.66	68.19	69.97	71.75

circ-SRBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	56.74	59.28	59.79	60.05	61.83	63.86	64.37	66.15	65.64	66.41	58.26	59.03	61.83	63.10	64.37	65.64	67.17	67.93	59.28	63.86	63.10	63.61	62.84	65.90
$R_1 = 2$	54.70	56.99	58.01	60.05	63.10	63.61	65.39	66.15	67.93	69.46	53.94	56.99	62.34	64.37	65.39	66.92	68.19	68.19	53.94	58.52	61.32	65.39	66.15	65.90
$R_1 = 3$	53.68	55.47	55.97	57.76	63.35	64.12	64.63	65.39	67.43	67.93	52.67	54.96	61.57	64.63	65.39	66.92	67.43	69.72	53.18	56.99	62.59	64.63	65.39	66.66
$R_1 = 4$	53.18	55.21	55.97	60.81	62.34	64.37	65.13	66.66	67.17	68.19	51.14	55.47	62.59	64.88	65.39	66.15	68.44	67.43	51.14	59.54	61.57	64.63	66.66	67.43
$R_1 = 5$	53.18	56.23	59.79	61.32	62.08	63.86	64.37	67.43	68.70	68.19	52.92	58.52	62.59	64.12	66.15	67.43	67.43	68.44	52.41	60.81	62.34	64.37	67.43	67.93
$R_1 = 6$	55.47	58.77	60.05	61.57	63.35	65.39	65.64	67.43	67.93	68.44	53.43	60.55	61.83	64.63	66.41	67.93	68.19	68.70	54.96	58.77	63.10	65.39	67.68	68.70

TABLEAU 6.6 – Taux d'identification (%) du descripteur circ-SMBP sur la base Trunk12 avec différentes représentations de la macrostructure.

circ-SMBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$							$\tau = 3$						
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	57.50	53.43	58.01	61.32	62.84	62.84	66.15	67.17	68.19	71.50	50.89	54.96	62.59	66.41	67.68	74.04	72.26	74.80	49.61	57.25	64.37	69.72	71.24	75.31
$R_1 = 2$	49.10	57.76	58.77	62.59	63.61	67.93	68.19	69.97	73.02	74.55	44.27	61.83	61.83	68.44	72.51	74.55	74.55	77.60	45.29	61.32	67.93	73.02	73.28	76.08
$R_1 = 3$	46.81	58.01	57.25	66.15	65.39	69.46	72.01	73.53	75.31	76.33	42.23	63.35	64.12	69.46	73.79	75.31	77.35	77.86	44.27	66.41	71.24	74.55	73.28	77.35
$R_1 = 4$	47.58	61.83	59.28	65.39	69.46	72.01	73.53	74.55	76.59	76.84	45.03	66.15	67.68	72.77	75.06	76.84	77.60	76.33	48.60	68.44	73.28	75.06	75.57	76.33
$R_1 = 5$	48.09	62.08	60.05	69.72	67.68	73.28	75.31	75.82	76.84	77.35	46.56	66.66	69.46	72.77	73.53	75.82	76.33	75.82	44.27	68.70	72.26	74.55	76.33	77.09
$R_1 = 6$	52.67	65.90	62.84	69.46	68.95	72.01	72.77	73.02	75.31	75.82	47.07	66.92	69.72	73.02	73.02	77.09	76.08	75.31	45.29	71.24	73.02	74.04	75.82	77.09

circ-SMBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$							$\tau = 3$						
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	56.99	56.48	59.54	59.79	61.83	61.83	63.61	66.41	67.43	70.99	50.12	54.19	56.48	62.59	66.92	69.46	70.99	72.51	47.83	54.19	60.30	67.68	70.73	70.22
$R_1 = 2$	50.38	59.54	59.28	59.03	60.55	65.39	64.88	67.93	72.01	72.01	50.89	57.50	58.26	63.86	67.17	71.24	72.26	71.50	49.10	57.50	64.12	68.95	71.75	72.26
$R_1 = 3$	49.10	58.26	58.01	62.59	60.05	65.90	66.15	70.22	70.73	72.01	50.38	57.76	60.05	64.63	70.48	72.26	72.26	71.50	46.05	58.26	65.13	69.46	72.26	71.50
$R_1 = 4$	51.39	60.05	55.97	62.59	62.59	67.17	67.68	70.22	72.51	73.79	46.56	59.79	62.08	64.63	70.99	73.28	72.01	73.53	46.31	61.32	66.66	71.50	71.50	71.50
$R_1 = 5$	49.61	63.10	60.30	66.92	65.13	67.43	69.72	72.77	73.28	73.02	50.89	61.57	62.59	67.93	71.24	73.28	73.28	73.02	46.56	62.08	67.68	71.75	73.79	71.75
$R_1 = 6$	50.12	65.39	62.59	65.39	65.13	69.21	72.01	72.51	73.53	74.04	49.36	63.86	63.10	69.46	72.26	73.28	73.79	72.01	48.85	63.35	67.17	71.50	74.04	72.26

6.2.1.1.2 Conception de l'espace d'échelle

Dans la précédente étude, on a noté que les scores des descripteurs statistiques ont tendance à augmenter en fonction de la taille de la macrostructure. Cependant, cette représentation génère des descripteurs avec une dimension considérable. Afin de remédier à cela, on fait varier différentes représentations de la macrostructure pour trouver la représentation optimale qui permet d'obtenir les meilleures performances d'identification des différentes espèces d'écorce.

Les Tableaux 6.5 et 6.6 présentent respectivement les taux d'identification pour les descripteurs SRBP et SMBP pour toutes les représentations possibles de la macrostructure en variant les paramètres de résolution spatiale et du voisinage. Les valeurs de \mathcal{L} sont fixées pour des tailles approximatives du support spatial qui englobe la macrostructure pour les différentes distances radiales τ .

Commençons par la localité du voisinage, paramétrée par les valeurs de $R_1 \in \llbracket 1, 6 \rrbracket$, on remarque à partir des Tableaux 6.5 et 6.6 que les taux d'identification augmentent graduellement en fonction de la valeur de R_1 , et ceux pour la majorité des valeurs de densité de résolution spatiale et du voisinage multi-échelles.

Ceci peut être expliqué par le fait que pour un voisinage très local autour du pixel central x_c , il n'y a pas de variation significative de la distribution d'intensité, i.e., les pixels voisins ont des valeurs d'intensité très similaires à celle du pixel central, et donc la distribution d'intensité dans ce voisinage n'est pas informative. Par contre, un voisinage distant à un rayon R_1 du pixel central est plus informatif et peut représenter la macrostructure de l'espèce d'écorce. Ainsi dit, on peut déduire que les pixels lointains contribuent au mieux à encoder une information de texture pertinente pour représenter la structure d'écorce.

Par rapport à la distance radiale τ , on peut constater à partir des résultats présentés dans les Tableaux 6.5 que modifier la représentation de l'espace d'échelle selon les valeurs de τ permet d'améliorer les scores d'identification pour un nombre réduit d'échelles, et ainsi générer un descripteur de dimension réduite.

Par exemple, pour le descripteur SRBP (cf. Tableau 6.5), on obtient 70.22% avec $\mathcal{L} = 11$, lorsque $\tau = 2$ est utilisé, on trouve 70.73% avec moins d'échelles ($\mathcal{L} = 9$), ce score est d'avantage amélioré (71.75%), avec $\tau = 3$ et $\mathcal{L} = 7$.

Ceci est aussi vérifié pour le descripteur SMBP (cf. Tableau 6.6), on obtient un score de 77.35% pour $\mathcal{L} = 11$, et 77.86% pour $\mathcal{L} = 9$ avec $\tau = 2$, ou encore avec moins d'échelles en $\tau = 3$ avec $\mathcal{L} = 7$ qui retourne 77.35%.

En dernier lieu, on examine l'influence de la densité du voisinage qui est aussi exprimée dans ces tableaux selon les deux schémas *AllN* et *OneN*. On remarque qu'une représentation dense du voisinage multi-échelles en *AllN* donne les meilleurs scores pour les descripteurs SRBP et SMBP. Une raison pour expliquer l'effet d'un voisinage dense sur les performances de nos descripteurs est essentiellement liée à la phase de la représentation et résumé statistique de la distribution d'intensité dans la macrostructure. En effet, avec le schéma *AllN*, un grand nombre de points voisins est pris en considération pour représenter la distribution d'intensité. Par conséquent, les représentations statistiques sont plus informatives que lorsqu'un faible nombre de points voisins est utilisé, comme le cas du schéma *OneN*.

Jusqu'à présent, nous avons évalué la représentation circulaire de la macrostructure. Dans la

6. Résultats expérimentaux 6.2. Évaluation des descripteurs de macro-motifs binaires statistiques

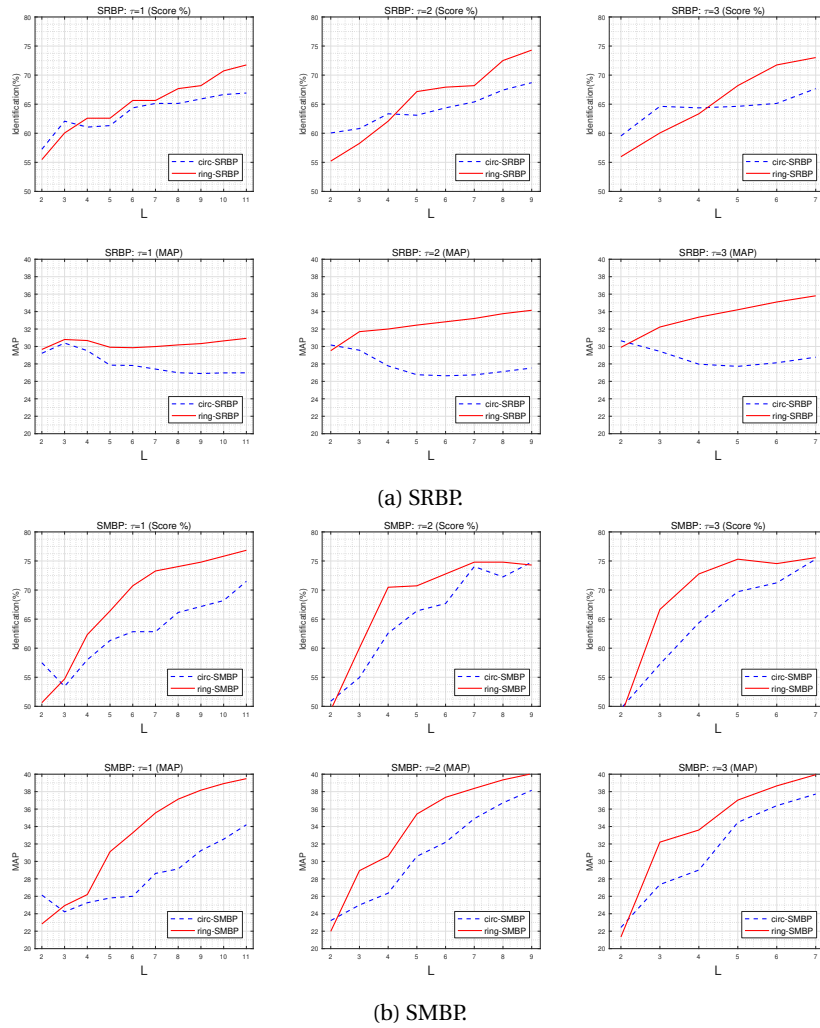


FIGURE 6.7 – Comparaison des scores d’identification (%) et des valeurs de MAP pour les descripteurs SRBP et SMBP avec les configurations *circ* et *ring* sur la base Trunk12.

TABLEAU 6.7 – Taux d’identification (%) des descripteurs ring-SRBP et ring-SMBP sur la base Trunk12 avec différentes représentations de la macrostructure.

ring-SRBP											OneN										
<i>AllN</i>											<i>OneN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	55.47	60.05	62.59	62.59	65.64	65.64	67.68	68.19	70.73	71.75	$\tau = 2$	55.47	59.79	61.57	64.37	65.13	67.68	67.17	67.17	68.19	71.24
$\tau = 3$	55.21	58.26	62.08	67.17	67.93	68.19	72.51	74.30	75.31	77.60	$\tau = 3$	54.70	59.28	62.84	67.93	73.79	66.92	70.22	70.73	72.77	72.77
$\tau = 4$	55.97	60.05	63.35	68.19	71.75	73.02	75.57	77.60	77.35	77.60	$\tau = 4$	55.21	59.54	62.34	67.68	70.48	70.99	72.26	73.28	73.79	73.02
$\tau = 5$	57.50	59.54	66.41	70.22	73.02	75.31	75.31	76.08	77.09	76.33	$\tau = 5$	56.74	60.55	66.15	70.22	71.24	72.51	72.77	75.57	75.57	75.57
ring-SMBP											OneN										
<i>AllN</i>											<i>OneN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	50.63	54.70	62.34	66.41	70.73	73.28	74.04	74.80	75.82	76.84	$\tau = 2$	49.36	56.48	59.28	63.10	67.43	70.48	71.75	72.26	73.53	74.55
$\tau = 3$	49.36	60.05	70.48	70.73	72.77	74.80	74.80	74.30	76.08	76.33	$\tau = 3$	52.67	60.55	67.68	71.24	73.28	73.53	72.77	75.06	74.80	75.82
$\tau = 4$	48.34	66.66	72.77	75.31	74.55	75.57	76.59	77.60	80.40	81.42	$\tau = 4$	48.60	62.08	70.99	73.53	72.77	73.28	74.55	74.55	76.84	78.62
$\tau = 5$	46.05	69.97	72.26	74.30	74.80	77.35	79.64	81.17	80.40	80.66	$\tau = 5$	50.63	64.63	71.50	69.97	73.28	75.31	75.82	79.88	79.38	78.37

suite, nous examinons une représentation en anneaux de la macrostructure pour les descripteurs radiaux et profonds.

La Figure 6.7 représente les scores d’identification et les valeurs de MAP réalisés sur la base Trunk12 avec les configurations équivalentes en *circ* et *ring* pour les différentes valeurs de τ . On

TABLEAU 6.8 – Configurations retenues pour les différents descripteurs statistiques sur la base Trunk12.

Descripteur	Configuration	1-NN-LOO±std	MAP	Dim	Descripteur	Configuration	1-NN-LOO±std	MAP	Dim
circ-SRBP	{ $\mathcal{L} = 7, R_1 = 6, \tau = 3, AllN$ }	71.75±3.23	36.35	640	ring-SRBP	{ $\mathcal{L} = 11, \tau = 7, AllN$ }	79.64±3.30	42.35	10240
circ-SRBP	{ $\mathcal{L} = 9, R_1 = 3, \tau = 2, OneN$ }	69.72±3.00	31.98	2560	ring-SRBP	{ $\mathcal{L} = 11, \tau = 7, OneN$ }	77.35±3.25	40.98	10240
Descripteur	Configuration	1-NN-LOO±std	MAP	Dim	Descripteur	Configuration	1-NN-LOO±std	MAP	Dim
circ-SMBP	{ $\mathcal{L} = 9, R_1 = 3, \tau = 2, AllN$ }	77.86±2.43	39.98	2560	ring-SMBP	{ $\mathcal{L} = 9, \tau = 5, AllN$ }	81.17±3.09	43.54	2560
circ-SMBP	{ $\mathcal{L} = 6, R_1 = 6, \tau = 3, OneN$ }	74.04±3.55	32.48	320	ring-SMBP	{ $\mathcal{L} = 9, \tau = 5, OneN$ }	79.88±3.03	43.29	2560

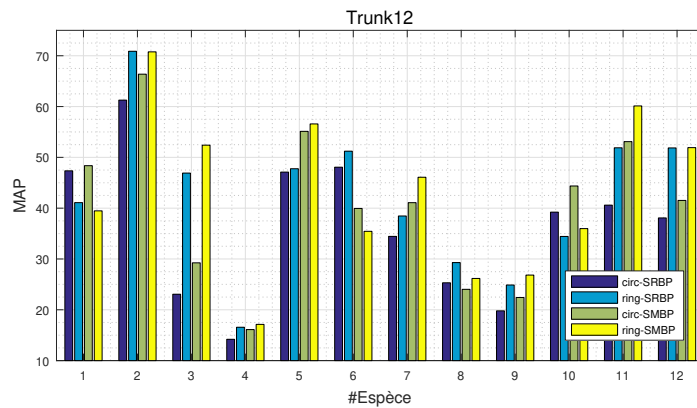


FIGURE 6.8 – Valeurs MAP par espèce des meilleures configurations des différents descripteurs statistiques, avec *AllN* sur la base Trunk12.

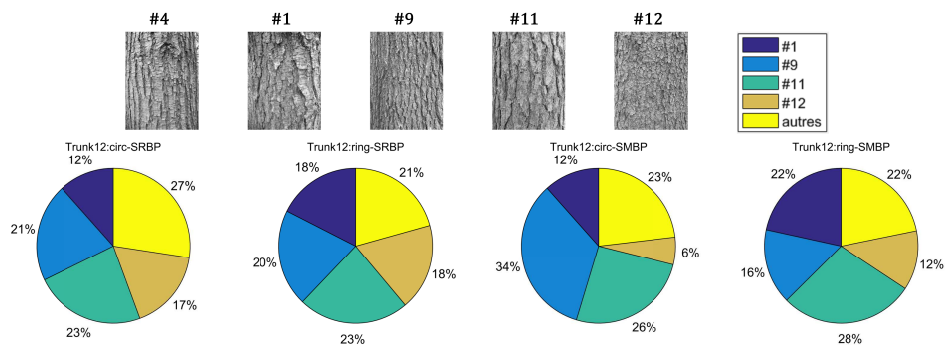


FIGURE 6.9 – Taux d’erreur de confusion de l’espèce #4 avec les espèces #1, #9, #11, #12 de la base Trunk12.

peut voir à partir de la Figure 6.7 que l’évolution du score d’identification avec une macrostructure *ring* dépasse celle d’une macrostructure *circ* en fonction de la taille de l’espace d’échelle.

Suivant une étude expérimentale similaire à celle réalisée précédemment avec une macrostructure circulaire *circ*, on commence par une étude évaluative des différents paramètres qui régularisent la représentation de la macrostructure en anneaux *ring*. Le Tableau 6.7 montre les taux d’identification en fonction de la taille de l’espace d’échelle \mathcal{L} pour différentes représentations τ de la largeur de l’anneau.

À partir du Tableau 6.7, on remarque que les scores d’identification pour le descripteur radial et profond augmentent en fonction de la taille de l’espace d’échelle \mathcal{L} , et la largeur de l’anneau régularisée par le paramètre τ . Cependant, plus la valeur de τ augmente plus la taille finale du descripteur baisse. Par exemple, pour le descripteur SRBP en *AllN* on trouve 71.75% avec $\mathcal{L} = 11$ et $\tau = 2$, et ce score s’améliore à 77.60% avec $\mathcal{L} = 9$ et $\tau = 4$. Pour ce qui est de la représentation des points voisins dans la macrostructure, les meilleurs scores sont retenus pour le schéma *AllN*.

Le Tableau 6.8 résume les meilleures configurations retenues pour chacun des descripteurs

statistiques avec les différentes représentations de l'espace d'échelle. On peut voir que globalement, la représentation de la macrostructure en *ring* améliore les performances d'une représentation en *circ*. Ceci est illustré par la Figure 6.8 où on affiche les valeurs de MAP par espèce pour chacun des descripteurs SRBP et SMBP en *circ* et en *ring*. Les configurations utilisées sont celles du Tableau 6.8 pour *AllN*.

À partir de la Figure 6.8, on remarque que parmi toutes les classes d'écorce de la base Trunk12, l'espèce #4 est la plus faiblement identifiée avec tous les descripteurs et les représentations de la macrostructure. En effet, les écorces de cette classe présentent une similarité d'apparence avec les espèces #1, #9, #11 et #12 comme le montre la Figure 6.9. Par ailleurs, cette similarité introduit une confusion lors de l'identification, et donc des taux d'erreur considérables.

6.2.1.1.3 Combinaisons des descripteurs statistiques

La représentation de la texture d'écorce par la combinaison de tous les descripteurs statistiques permet d'obtenir des scores d'identification élevés comparée à une représentation à base de descripteurs statistiques singuliers. Cependant, il sera intéressant, en terme de la dimensionnalité de la représentation, de pouvoir dériver une représentation de la texture d'écorce formée à base d'une combinaison moindre de descripteurs statistiques. De ce fait, on analyse l'apport des différentes représentations basées sur les combinaisons de deux, trois, et quatre descripteurs statistiques. Dans la Figures 6.10 on représente les performances des combinaisons possibles à base de deux descripteurs statistiques, où la couleur de chaque nœud indique le taux d'identification et la valeur MAP associée à chaque descripteur singulier SRBP_{*m*} (respectivement SMBP_{*m*}), et la largeur de l'arrête reliant deux nœuds représente le score d'identification et la valeur MAP de la combinaison de ces deux descripteurs.

À partir de la Figure 6.10, on remarque que :

- Les descripteurs formés par une combinaison de deux descripteurs améliorent d'avantage les performances des descripteurs individuels.
- Pour les deux encodages des descripteurs, les descripteurs statistiques de variance (Var), i.e., SRBP_{Var} et SMBP_{Var} retournent les meilleurs scores suivis par ceux de la moyenne (Mean). En effet, la représentation par la moyenne et la variance approxime la distribution d'intensité par le moment du premier et second ordre dans la macrostructure, et donc représente une information intrinsèque pertinente.
- Parmi les combinaisons deux à deux possibles, c'est SRBP_{MeanVar} et SMBP_{MinMax} qui retournent les meilleures valeurs de MAP. Le descripteur SMBP_{MinMax} retourne un score intéressant avec une amélioration moyenne de +13.10% par rapport aux descripteurs SMBP_{Min} et SMBP_{Max}. En effet, l'encodage des variations des valeurs extrêmes d'intensité permet de représenter les changements de valeurs d'intensité locales dans la surface des écorces qui caractérisent leurs apparences (comme des nuances de gris pour certaines espèces comme l'espèce #2).

Par ailleurs, nous rapportons aussi les performances des trois meilleures combinaisons de descripteurs statistiques à base de trois et quatre descripteurs dans le Tableau 6.9, où on peut voir que, globalement, pour les deux encodages proposés, c'est la combinaison des cinq descripteurs statis-

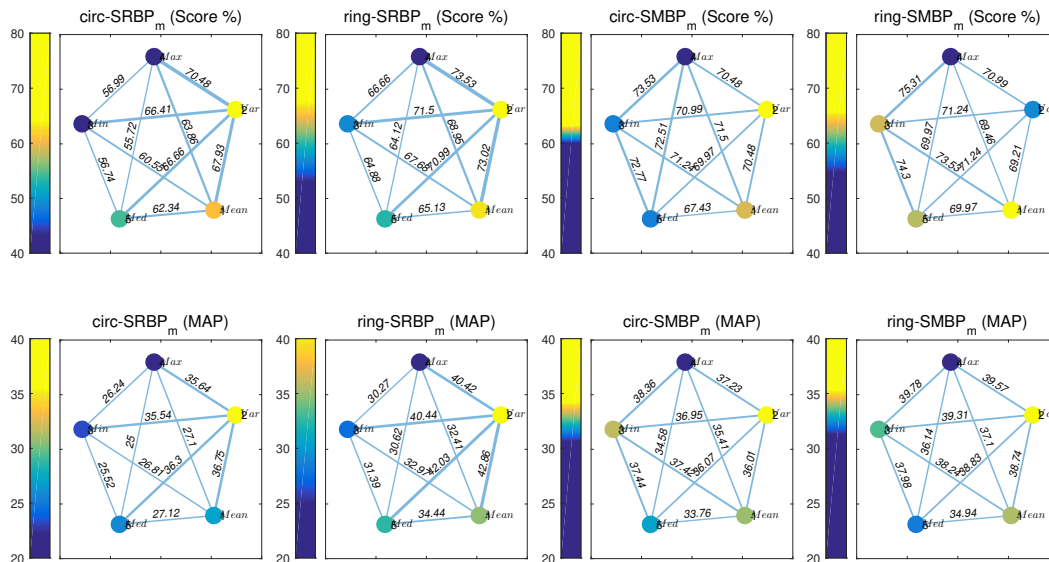


FIGURE 6.10 – Taux d'identification (%) et valeurs MAP des combinaisons formées de deux descripteurs statistiques sur la base Trunk12.

TABLEAU 6.9 – Taux d'identification (%) et valeurs MAP des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base Trunk12.

circ-SRBP ($\mathcal{L} = 7$)				ring-SRBP ($\mathcal{L} = 11$)			
Descripteur	1-NN-LOO \pm std	MAP	Dim	Descripteur	1-NN-LOO \pm std	MAP	Dim
SRBP _{MinMaxVar}	70.73 \pm 2.67	35.84	384	SRBP _{MaxMedVar}	75.31 \pm 3.58	41.50	6144
SRBP _{MaxMedVar}	71.24 \pm 2.93	35.98	384	SRBP _{MaxMeanVar}	76.59 \pm 3.33	42.10	6144
SRBP _{MaxMeanVar}	72.01\pm2.92	36.47	384	SRBP _{MinMeanVar}	76.59 \pm 3.56	42.39	6144
SRBP _{MinMaxMedVar}	70.48 \pm 2.80	36.05	512	SRBP _{MaxMeanMedVar}	77.35 \pm 3.34	42.47	8192
SRBP _{MinMaxMeanVar}	71.00 \pm 3.22	36.31	512	SRBP _{MinMeanMedVar}	77.86 \pm 2.78	42.83	8192
SRBP _{MaxMeanMedVar}	71.00 \pm 2.89	36.48	512	SRBP _{MinMaxMeanVar}	79.13 \pm 3.22	42.06	8192
SRBP	71.75 \pm 3.23	36.35	640	SRBP	79.64\pm3.30	42.35	10240

circ-SMBP ($\mathcal{L} = 9$)				ring-SMBP ($\mathcal{L} = 9$)			
Descripteur	1-NN-LOO \pm std	MAP	Dim	Descripteur	1-NN-LOO \pm std	MAP	Dim
SMBP _{MinMedVar}	75.31 \pm 4.46	38.46	1536	SMBP _{MinMaxMed}	78.88 \pm 3.74	41.64	1536
SMBP _{MinMaxMean}	76.08 \pm 3.92	39.48	1536	SMBP _{MinMaxMean}	79.38 \pm 3.41	41.63	1536
SMBP _{MinMaxMed}	76.33 \pm 3.59	39.71	1536	SMBP _{MinMaxVar}	79.38 \pm 2.88	43.39	1536
SMBP _{MinMaxMeanVar}	76.08 \pm 3.98	39.88	2048	SMBP _{MinMaxMeanVar}	78.88 \pm 3.57	43.58	2048
SMBP _{MinMaxMedVar}	76.59 \pm 3.94	40.07	2048	SMBP _{MinMaxMedVar}	80.15 \pm 3.33	43.75	2048
SMBP _{MinMaxMeanMed}	77.60 \pm 3.81	39.73	2048	SMBP _{MinMaxMeanMed}	80.66 \pm 2.90	41.81	2048
SMBP	77.86\pm4.22	39.98	2560	SMBP	81.17\pm3.09	43.54	2560

tiques qui obtient les scores les plus élevés. De plus, pour les combinaisons à quatre descripteurs, on trouve celle formée des descripteurs à bases des mesures statistiques Mean, Var, Min et Max qui vient en seconde position après celle de tous les descripteurs. Ceci montre la complémentarité de nos descripteurs à représenter la surface des différentes écorces en plus de réduire la taille finale du descripteur par un facteur d'échelle.

6.2.1.2 Comparaison avec l'état de l'art

Le Tableau 6.10 présente les résultats d'identification des espèces d'écorce, exprimés par le score d'identification de 1-NN-LOO et les valeurs de MAP. En observant le Tableau 6.10, on peut voir que les meilleurs résultats sur la base Trunk12 sont obtenus par 81 descripteurs, notamment

TABLEAU 6.10 – Résultats de comparaison avec l'état de l'art sur la base Trunk12. les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].

Ref	Descripteur	1-NN-LOO±std	MAP	Dim
[WAN et collab., 2004]	GLCM	44.52±4.90	27.43	20
[HUANG et collab., 2006d]	Gabor	48.09±1.57	23.42	48
[HUANG et collab., 2006b]	DWT	31.20±3.42	22.77	19
[BERTRAND et collab., 2017]	Gw _s [*]	39.90 [*]	—	121
[OJALA et collab., 2002]	LBP	48.85±3.25	26.47	256
[OJALA et collab., 2002]	MResLBP	56.74±5.77	26.24	54
[MÄENPÄÄ et PIETIKÄINEN, 2003]	LBPf	60.81±3.00	27.97	30
[ŠULC et MATAS, 2013]	MSLBP	65.13±2.51	27.63	180
[QIAN et collab., 2011]	PLBP	66.92±3.38	28.63	162
[SIXTA, 2011]	MBLBP	54.45±4.12	27.63	100
[WOLF et collab., 2008]	TPLBP	52.41±2.93	24.13	256
[WOLF et collab., 2008]	FPLBP	37.40±2.54	25.18	16
[WANG et collab., 2013]	LNIRP/LBP/DCI-PTP	74.55±2.08	33.32	600
[LIU et collab., 2016b]	MRELBP	75.06±2.83	34.93	800
[GUO et collab., 2010b]	LBPV	53.94±5.77	23.09	54
[NGUYEN et collab., 2016]	MC-SBP2	62.08±1.96	32.73	4000
[RATAJCZAK et collab., 2019]	LS-LCoLBP [*]	74.60 [*]	—	90
[QI et collab., 2013]	MSJLBP	60.81±4.18	27.98	3540
[QI et collab., 2014]	PRICoLBP	67.17±2.81	29.12	4000
[LIU et collab., 2012b]	NI/RD/CI	50.38±2.30	20.76	2200
[LIU et collab., 2016a]	ELBP ₁₃₄₆	61.06±3.33	24.36	80
[LIU et collab., 2016a]	ELBP ₁₃₄₅₆	61.57±3.61	24.68	100
[LIU et collab., 2016a]	ELBP ₁₂₃₄₅₆	61.32±3.46	25.52	120
[GUO et collab., 2010a]	CLBP-S/M/C	59.03±1.88	31.24	2200
[ZHAO et collab., 2012b]	CLBC-S/M/C	67.68±1.19	32.14	1990
[LIU et collab., 2014]	BRINT-CS-CM	69.21±1.66	32.00	360
[proposé]	circ-SRBP _{MaxMeanVar}	72.01±2.92	36.47	384
[proposé]	circ-SRBP	71.75±3.23	36.35	640
[proposé]	SRBP _{MinMaxMeanVar}	79.13±3.22	42.06	8192
[proposé]	SRBP	79.64±3.30	42.35	10240
[proposé]	circ-SMBP _{MinMaxMeanMed}	77.60±3.81	39.73	2048
[proposé]	circ-SMBP	77.86±4.22	39.98	2560
[proposé]	ring-SMBP _{MinMaxMeanMed}	80.66±2.90	41.81	2048
[proposé]	ring-SMBP	81.17±3.09	43.54	2560

le descripteur SMBP avec une représentation en anneaux, ring-SMBP, qui retourne un score de 81.17% et une MAP de 43.54.

Lors de cette comparaison, nous avons choisi de se comparer avec des méthodes classiques, comme les GLCM, Gabor et les ondelettes DWT. Il est clair que ces approches ne sont plus suffisantes pour encoder des attributs pertinents de la surface des écorces, et sont dépassées par les variantes LBP, notamment les variantes multi-échelles, comme la LNIRP/LBP/DCI-PTP[WANG et collab., 2013] basées sur l'utilisation de patchs moyens autour des pixels voisins, et la MRELBP[LIU et collab., 2016b] basée sur l'utilisation des patchs médians.

Bien que les résultats des descripteurs SRBP restent légèrement faibles par rapport à ceux obtenus par SMBP, ils ont de meilleures performances par rapport aux variantes qui encodent les variations inter-échelles comme PRICoLBP[QI et collab., 2014], comme pas exemple, pour circ-SRBP_{MaxMeanVar} avec un score de 72.01%, et une MAP de 36.47 contre un score de 67.17% et une MAP de 29.12 pour PRICoLBP.

La catégorie qui se rapproche le plus de notre approche est celle des descripteurs statistiques comme MC-SBP2[NGUYEN et collab., 2016] et LS-LCoLBP[RATAJCZAK et collab., 2019], qui obtient

le second meilleur score sur cette base avec 74.60%. Ceci est dû au fait que LS-LCoLBP combine plusieurs encodages de gradients locaux autour du pixel.

Au final, les résultats obtenus sur la base Trunk12 montrent la pertinence des descripteurs proposés, en particulier les descripteurs à encodage profond SMBP en *ring*. Ces résultats restent à confirmer sur les autres bases d'écorce qui contiennent plus d'images d'écorce photographiées dans des conditions plus contraignantes.

6.2.2 Résultats sur la base AFF

6.2.2.1 Évaluation des paramètres

6.2.2.1.1 Conception de l'espace d'échelle

Afin de trouver la représentation optimale de la macrostructure, on teste différentes valeurs des paramètres qui régularisent sa taille au niveau de la résolution spatiale et du voisinage. Les scores d'identification des différentes configurations résultantes sont présentés dans les Tableaux 6.11 et 6.12. Les descripteurs globaux SRBP, et SMBP sont utilisés pour la description finale des images d'écorce. À partir de ces tableaux, on rapporte les remarques suivantes :

On commence par l'emplacement spatial du voisinage par rapport au pixel central, paramétré par les valeurs de R_1 , où on peut voir à partir des tableaux que les grandes valeurs de R_1 contribuent à mieux améliorer le score de l'identification des espèces d'écorce. Des résultats similaires étaient notés pour la base Trunk12 du fait que les pixels distants sont plus informatifs que les pixels très locaux.

Pour ce qui est de la densité en résolution spatiale et du voisinage, on peut voir à partir des Tableaux 6.11 et 6.12 que les meilleurs scores sont obtenus pour $\tau = 2$ et $AllN$ pour les descripteurs radiaux SRBP avec 76.90% et profonds SMBP avec 81.38%.

La Figure 6.14 montre les courbes des scores d'identification et les valeurs de MAP par rapport aux niveaux d'échelle \mathcal{L} pour une macrostructure *circ* et son équivalent en anneaux *ring*. On peut voir qu'une représentation en *ring* améliore légèrement les performances. De ce fait, on examine la configuration optimale pour une description en SRBP et SMBP avec une macrostructure *ring*. Le Tableau 6.13 montre l'évolution des performances en fonction des valeurs d'échelle et de la largeur d'anneau. Une fois de plus, on trouve que plus la largeur de l'anneau augmente, le score d'identification augmente. Au final, on résume les meilleures configurations retenues pour les deux encodages proposés dans le Tableau 6.14.

6.2.2.1.2 Combinaisons des descripteurs statistiques

Dans le but d'obtenir une représentation performante des structures d'écorce formée à base d'une combinaison d'un faible nombre de descripteurs singuliers, on présente les scores d'identification des différentes représentations de texture d'écorce à base de deux, trois, et quatre descripteurs statistiques, pour chacun des descripteurs $SRBP_m$ et $SMBP_m$. La Figure 6.12 montre l'apport des descripteurs seuls et combinés deux à deux, et le Tableau 6.15 présente les scores et les MAP des combinaisons formées de trois, quatre et tous les descripteurs statistiques pour les encodages radiaux et profonds.

TABLEAU 6.11 – Taux d'identification (%) du descripteur circ-SRBP sur la base AFF avec différentes représentations de la macrostructure.

circ-SRBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$							$\tau = 3$						
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	43.40	46.95	51.01	53.89	55.24	56.59	59.56	62.35	63.11	63.87	44.67	53.89	58.88	61.67	63.02	65.65	67.00	67.68	45.34	58.37	63.36	65.90	67.25	69.88
$R_1 = 2$	46.95	50.76	53.55	56.00	58.03	59.05	60.49	60.74	63.62	64.80	48.39	55.92	59.56	60.82	63.45	65.31	67.51	68.35	50.42	58.62	62.85	63.79	66.75	68.18
$R_1 = 3$	50.50	54.82	56.59	57.78	59.89	60.49	62.26	64.12	64.46	71.31	51.35	56.59	60.99	63.45	65.39	66.58	67.00	67.93	51.94	59.13	64.38	67.00	68.10	69.79
$R_1 = 4$	52.19	57.10	58.62	60.40	62.01	62.43	63.36	65.05	66.24	74.45	53.13	59.05	62.35	64.72	66.41	67.42	67.59	69.28	55.83	61.67	64.55	66.83	70.05	71.31
$R_1 = 5$	56.09	58.03	60.49	62.01	63.02	64.55	65.48	65.90	66.41	66.49	66.07	70.30	72.25	74.28	74.53	75.46	76.31	76.90	56.42	64.46	65.65	68.02	70.30	70.81
$R_1 = 6$	57.10	59.64	62.85	64.04	64.63	65.73	66.41	66.24	63.70	66.49	58.46	63.79	65.98	66.15	67.59	69.12	70.30	70.81	59.22	65.56	67.34	68.61	72.41	73.18
circ-SRBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$							$\tau = 3$						
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	41.20	44.92	49.57	51.43	56.17	57.52	59.89	58.46	59.22	60.06	42.38	50.33	54.65	56.93	57.95	61.33	62.01	63.11	43.23	55.66	59.22	61.84	64.04	64.89
$R_1 = 2$	46.44	50.16	52.87	53.97	55.66	57.02	58.62	59.30	60.23	61.08	47.63	52.11	54.82	57.44	60.06	61.08	62.18	63.53	48.39	55.58	59.22	62.01	63.87	65.39
$R_1 = 3$	48.64	52.87	55.49	56.34	58.96	61.08	61.16	61.67	61.59	62.09	50.42	56.51	58.88	61.08	62.69	62.77	63.53	64.29	50.08	58.03	63.19	64.80	65.22	65.82
$R_1 = 4$	51.43	54.48	59.22	60.06	60.91	63.11	63.36	65.22	65.65	66.07	50.59	57.69	60.40	62.85	63.19	63.70	64.38	64.80	50.93	60.15	63.36	64.89	65.65	66.15
$R_1 = 5$	52.96	56.68	59.72	61.59	63.02	63.70	64.38	65.22	65.73	65.05	53.38	60.40	62.52	63.28	62.85	63.87	65.73	65.05	54.31	62.09	64.63	65.73	66.58	66.41
$R_1 = 6$	52.70	57.69	58.54	59.72	61.33	61.50	62.52	62.43	62.69	63.79	53.46	61.16	64.29	64.04	64.97	65.56	65.48	66.24	56.34	63.53	65.14	66.32	66.66	67.00

TABLEAU 6.12 – Taux d'identification (%) du descripteur circ-SMBP sur la base AFF avec différentes représentations de la macrostructure.

circ-SMBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	42.13	50.16	60.57	63.19	64.89	67.34	69.03	72.25	71.57	72.25	39.93	57.86	64.55	69.20	70.55	71.65	71.06	72.33	45.76	62.09	67.00	68.78	71.06	72.08
$R_1 = 2$	43.82	56.26	63.70	67.25	69.28	71.57	73.26	72.16	72.67	73.77	47.96	63.87	70.55	72.16	71.65	72.50	73.60	74.11	49.49	64.55	69.54	70.89	72.92	73.60
$R_1 = 3$	47.54	62.85	67.68	69.71	72.08	73.18	73.68	73.35	74.45	77.66	50.16	65.65	71.82	71.48	73.35	73.60	74.87	74.95	52.62	67.42	72.58	73.35	74.02	76.22
$R_1 = 4$	47.20	65.56	69.96	71.40	73.85	73.09	73.94	74.78	74.78	75.04	51.60	67.76	73.68	73.18	75.12	75.97	75.88	75.38	57.95	69.45	73.68	73.60	74.70	75.54
$R_1 = 5$	48.13	66.41	69.28	71.15	72.33	73.68	75.04	75.12	75.04	74.78	62.09	74.45	77.58	80.03	79.86	80.54	81.13	81.38	56.42	69.62	74.36	73.94	75.46	75.71
$R_1 = 6$	52.70	66.92	70.13	70.98	73.85	74.53	76.14	75.29	72.84	75.97	52.96	69.03	74.19	74.19	75.12	75.80	75.54	75.46	54.56	70.13	74.78	74.95	77.66	75.71
circ-SMBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	42.89	53.13	57.95	61.33	65.82	68.10	68.61	66.41	70.05	71.74	42.21	55.16	61.42	65.90	68.35	68.35	68.69	69.54	44.33	58.20	65.82	67.42	67.68	67.59
$R_1 = 2$	43.99	54.14	60.32	61.75	64.63	67.51	69.37	71.40	71.57	71.57	47.03	58.20	66.58	67.85	69.54	69.45	70.47	70.89	46.02	61.92	67.76	68.27	67.93	69.71
$R_1 = 3$	46.44	57.69	62.43	64.80	68.10	69.28	70.55	71.82	71.91	71.40	47.20	61.67	68.10	69.20	69.96	69.62	71.23	71.23	46.27	62.69	69.20	68.69	69.12	71.23
$R_1 = 4$	42.04	60.23	63.87	66.75	70.05	70.64	72.08	72.50	71.48	71.57	48.90	62.09	68.02	69.03	71.06	70.89	71.31	72.33	53.04	64.04	69.54	69.03	69.20	70.64
$R_1 = 5$	46.70	61.42	64.72	67.59	69.71	71.06	71.65	71.65	72.25	73.01	57.78	63.11	68.61	69.79	70.38	70.81	71.31	71.74	54.14	64.38	69.71	68.78	70.98	71.57
$R_1 = 6$	57.27	61.33	65.22	67.42	68.95	69.79	70.89	71.74	70.89	70.98	53.13	63.19	67.85	68.44	69.28	70.05	70.47	71.91	53.38	65.48	68.35	69.12	71.57	71.74

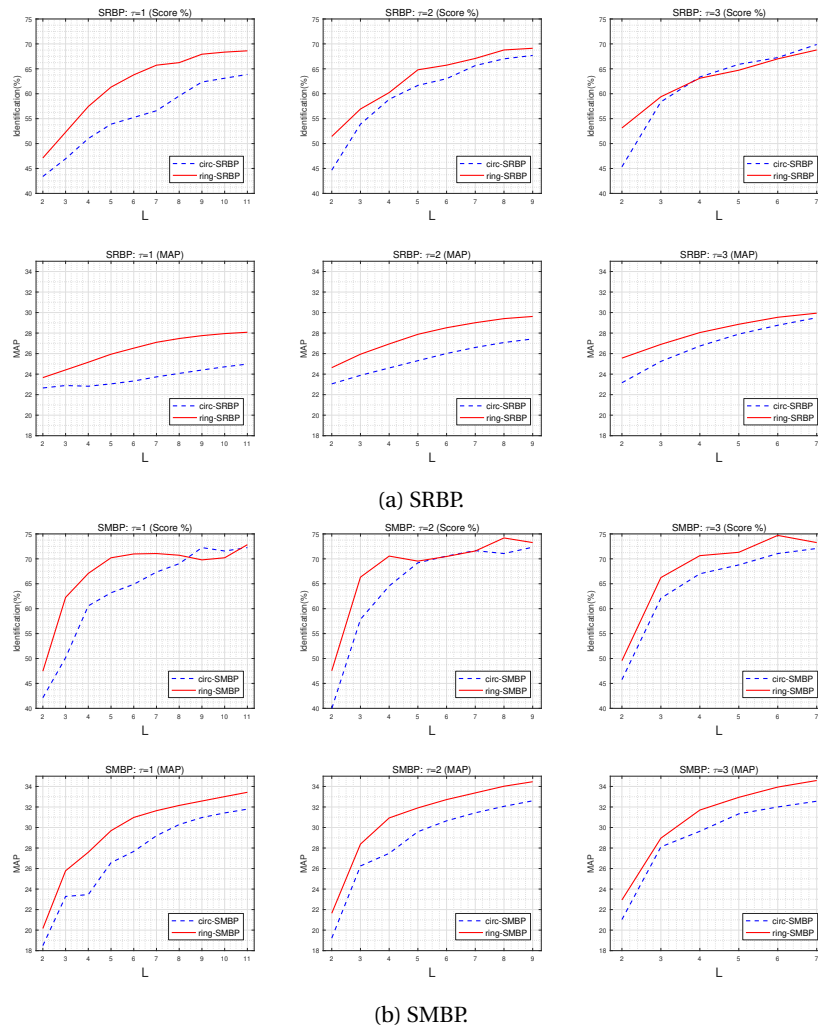


FIGURE 6.11 – Comparaison des scores d’identification (%) et des valeurs de MAP pour les descripteurs SRBP et SMBP avec les configurations *circ* et *ring* sur la base AFF.

À partir de la Figure 6.12 et du Tableau 6.15 on peut voir que plus on combine des descripteurs statistiques, plus les performances augmentent, d’où la complémentarité des représentations. Parmi les combinaisons des descripteurs statistiques, c’est ceux basés sur les mesures statistiques Mean, Var, Min et Max qui retournent les meilleures performances, et ceux pour les différents encodages et représentations *circ* et *ring*.

Ces comportements de nos descripteurs sont similaires à ceux trouvés sur la base Trunk12, et ceux malgré le fait que la base AFF contient trois fois plus d’images que la base Trunk12. De plus, les images de AFF sont prises dans des conditions non-contrôlées, d’où le fait que c’est une base défiante. Ceci montre que notre modélisation en macro-motifs statistiques est pertinente pour la représentation des textures d’écorce, reste à vérifier leurs performances contre celles des méthodes de l’état de l’art.

6.2.2.2 Comparaison avec l’état de l’art

Les résultats comparatifs pour l’identification des espèces végétales sur la base AFF sont présentés dans le Tableau 6.16. À partir du Tableau 6.16 on peut voir que :

TABLEAU 6.13 – Taux d’identification (%) des descripteurs ring-SRBP et ring-SMBP sur la base AFF avec différentes représentations de la macrostructure.

ring-SRBP											
AllN						OneN					
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	
$\tau = 2$	47.12	52.28	57.44	61.33	63.79	65.73	66.24	67.93	68.35	68.61	
$\tau = 3$	51.43	56.93	60.23	64.80	65.73	67.08	68.78	69.12	70.38	71.31	
$\tau = 4$	53.13	59.39	63.11	64.72	67.00	68.78	69.96	72.33	72.16	72.75	
$\tau = 5$	56.59	60.82	63.95	66.58	68.69	70.38	71.99	72.41	73.60	74.61	

ring-SMBP											
AllN						OneN					
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	
$\tau = 2$	47.46	62.26	67.08	70.21	70.98	71.06	70.72	69.79	70.21	72.84	
$\tau = 3$	47.54	66.32	70.55	69.54	70.47	71.57	74.19	73.26	73.26	73.09	
$\tau = 4$	49.57	66.24	70.64	71.31	74.70	73.26	73.68	72.84	72.92	71.40	
$\tau = 5$	52.11	67.08	70.38	72.58	73.18	72.67	72.92	71.91	71.74	70.98	

TABLEAU 6.14 – Configurations retenues pour les différents descripteurs statistiques sur la base AFF.

Descripteur	Configuration	1-NN-LOO±std	MAP	Dim	Descripteur	Configuration	1-NN-LOO±std	MAP	Dim
circ-SRBP	{ $\mathcal{L} = 9, R_1 = 5, \tau = 2, AllN$ }	76.90±2.83	26.49	2560	ring-SRBP	{ $\mathcal{L} = 10, \tau = 6, AllN$ }	74.78±0.76	32.03	5120
circ-SRBP	{ $\mathcal{L} = 7, R_1 = 6, \tau = 3, OneN$ }	67.00±1.38	27.65	640	ring-SRBP	{ $\mathcal{L} = 9, \tau = 7, OneN$ }	73.77±0.88	31.52	2560

Descripteur	Configuration	1-NN-LOO±std	MAP	Dim	Descripteur	Configuration	1-NN-LOO±std	MAP	Dim
circ-SMBP	{ $\mathcal{L} = 9, R_1 = 5, \tau = 2, AllN$ }	81.38±2.23	31.50	2560	ring-SMBP	{ $\mathcal{L} = 6, \tau = 4, AllN$ }	74.70±0.86	33.94	320
circ-SMBP	{ $\mathcal{L} = 11, R_1 = 5, \tau = 1, OneN$ }	73.01±0.79	30.89	10240	ring-SMBP	{ $\mathcal{L} = 9, \tau = 4, OneN$ }	73.18±1.35	34.17	2560

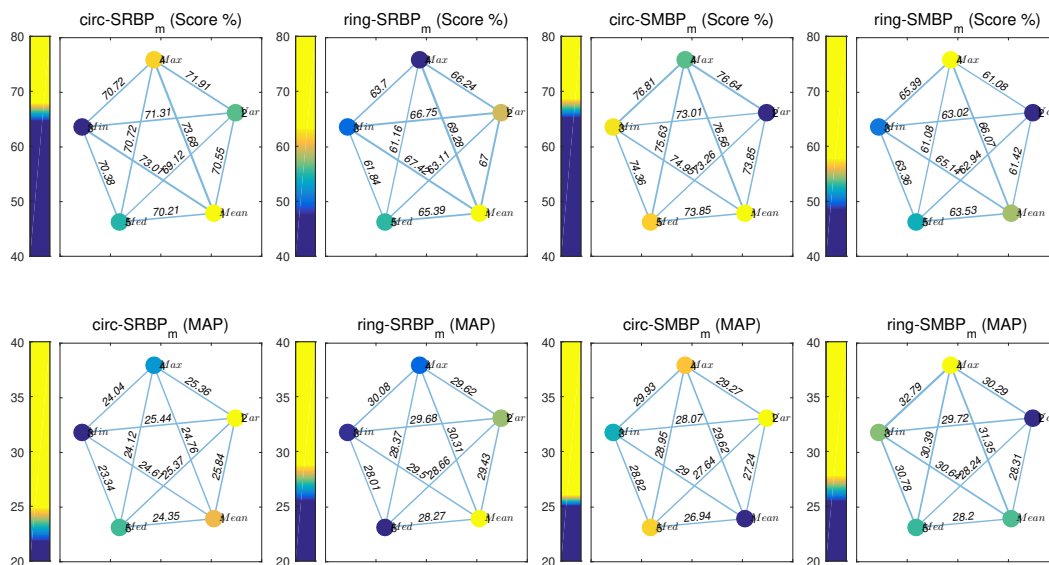


FIGURE 6.12 – Taux d’identification (%) et valeurs MAP des combinaisons formées de deux descripteurs statistiques sur la base AFF.

- Tous les descripteurs proposés ainsi que leurs combinaisons obtiennent les meilleurs scores d’identification, et généralement des valeurs de MAP des descripteurs circ-SRBP légèrement inférieures à la meilleure valeur obtenue par le descripteur MRELBP [LIU et collab., 2016b]. Le descripteur circ-SMBP retourne un score de 81.38% et une MAP de 31.5 qui dépassent significativement tous les scores et les MAP obtenus par les méthodes en comparaison sur la base AFF.
- Parmi les descripteurs proposés, les descripteurs radiaux, i.e., SRBP, et SRBP_{MeanVarMinMax}, retournent des scores relativement inférieurs à ceux obtenus par les descripteurs du macro-

TABLEAU 6.15 – Taux d’identification (%) et valeurs MAP des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base AFF.

circ-SRBP ($\mathcal{L} = 9$)				ring-SRBP ($\mathcal{L} = 10$)			
Descripteur	1-NN-LOO \pm std	MAP	Dim	Descripteur	1-NN-LOO \pm std	MAP	Dim
SRBP _{MinMeanMed}	74.11 \pm 3.27	24.70	1536	SRBP _{MinMaxVar}	70.38 \pm 0.80	31.15	3072
SRBP _{MinMaxVar}	74.87 \pm 2.86	25.90	1536	SRBP _{MaxMeanVar}	70.47 \pm 1.32	30.87	3072
SRBP _{MinMaxMean}	75.63 \pm 3.39	25.27	1536	SRBP _{MinMaxMean}	71.23 \pm 0.80	31.64	3072
SRBP _{MinMaxMeanMed}	75.54 \pm 3.47	25.29	2048	SRBP _{MinMaxMedVar}	72.08 \pm 0.82	31.60	4096
SRBP _{MinMaxMedVar}	76.14 \pm 2.79	26.10	2048	SRBP _{MinMaxMeanMed}	72.41 \pm 0.74	31.45	4096
SRBP _{MinMaxMeanVar}	76.81 \pm 2.76	26.42	2048	SRBP _{MinMaxMeanVar}	74.53 \pm 0.68	31.90	4096
SRBP	76.90\pm2.83	26.49	2560	SRBP	74.78\pm0.76	32.03	5120
circ-SMBP ($\mathcal{L} = 9$)				ring-SMBP ($\mathcal{L} = 6$)			
Descripteur	1-NN-LOO \pm std	MAP	Dim	Descripteur	1-NN-LOO \pm std	MAP	Dim
SMBP _{MinMaxVar}	79.61 \pm 2.33	30.86	1536	SMBP _{MinMaxMed}	69.45 \pm 0.94	34.05	192
SMBP _{MinMaxMed}	80.03 \pm 2.71	31.16	1536	SMBP _{MinMaxVar}	69.88 \pm 0.90	33.25	192
SMBP _{MinMaxMean}	80.45 \pm 2.32	31.61	1536	SMBP _{MinMaxMean}	71.23 \pm 1.00	34.10	192
SMBP _{MinMaxMedVar}	80.62 \pm 2.30	31.25	2048	SMBP _{MinMaxMeanMed}	72.41 \pm 1.09	34.42	256
SMBP _{MinMaxMeanMed}	80.79 \pm 3.34	31.87	2048	SMBP _{MinMaxMedVar}	73.26 \pm 0.72	33.75	256
SMBP _{MinMaxMeanVar}	81.21 \pm 2.01	31.43	2048	SMBP _{MinMaxMeanVar}	73.43 \pm 0.99	33.78	256
SMBP	81.38\pm2.23	31.50	2560	SMBP	74.70\pm0.86	33.94	320

motif profond. Cependant, nos descripteurs radiaux retournent des performances supérieures à tous les descripteurs LBP radiaux en comparaison, i.e., MSJLBP [QI et collab., 2013], PRICoLBP [QI et collab., 2014], et NI/RD/CI [LIU et collab., 2012b], et cela en terme de score et de MAP.

- Parmi les catégories des méthodes en comparaison, on trouve que les méthodes statistiques : LBPV [GUO et collab., 2010b], MC-SBP2 [NGUYEN et collab., 2016] ont une moyenne de scores supérieure aux autres catégories. Cependant, les descripteurs MRELBP [LIU et collab., 2016b] avec un patch médian, et LNIRP/LBP/DCI-PTP [WANG et collab., 2013] avec un patch moyen, obtiennent le second et le troisième meilleur score sur la base AFF. Ceci montre l’efficacité d’un encodage multi-échelles et celui de l’information statistique pour représenter la texture d’écorce. Par ailleurs, il est important de noter que ces deux descripteurs introduisent un descripteur du pixel central qui encode la variation du contraste global de l’image, tandis que nos descripteurs n’encodent pas d’autres variations que celles de la distribution entre les niveaux d’échelle de la macrostructure. Ceci montre l’efficacité de l’approche proposée.
- Toutefois, il est important de noter que toutes les combinaisons des différents descripteurs statistiques proposées et mentionnées dans le Tableau 6.15, notamment les descripteurs circ-SRBP_m, obtiennent des scores élevés par rapport à ceux de toutes les méthodes de l’état de l’art. De plus, ceci est vérifié aussi pour certaines combinaisons à base de deux descripteurs schématisées dans la Figure 6.12, comme le circ-SRBP_{MeanMax}, et le circ-SMBP_{MinMax} avec les scores et les MAP respectifs de 73.68%, 24.76 et 76.81%, 29.93. D’un côté ceci montre que la tâche d’identification d’objets naturels comme l’écorce est complexe et défiante, et de l’autre côté, ceci montre que nos descripteurs qui modélisent une approche statistique pour la description de la macrostructure de l’écorce sont plus pertinents pour distinguer les différentes structures des écorces.

TABLEAU 6.16 – Résultats de comparaison avec l'état de l'art sur la base AFF. les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].

Ref	Descripteur	1-NN-LOO±std	MAP	Dim
[WAN et collab., 2004]	GLCM	38.15±0.63	22.88	20
[HUANG et collab., 2006d]	Gabor	44.92±1.44	24.46	48
[HUANG et collab., 2006b]	DWT	42.55±1.26	24.78	19
[BERTRAND et collab., 2017]	Gw ₃ *	48.20*	—	121
[OJALA et collab., 2002]	LBP	47.63±2.72	21.10	256
[OJALA et collab., 2002]	MResLBP	52.70±1.76	20.47	54
[MÄENPÄÄ et PIETIKÄINEN, 2003]	LBPf	59.72±1.55	23.82	30
[ŠULC et MATAS, 2013]	MSLBP	63.28±1.68	25.40	180
[QIAN et collab., 2011]	PLBP	61.50±1.72	25.65	162
[SIXTA, 2011]	MBLBP	50.16±2.07	27.00	100
[WOLF et collab., 2008]	TPLBP	47.29±1.80	22.20	256
[WOLF et collab., 2008]	FPLBP	44.07±1.90	28.08	16
[WANG et collab., 2013]	LNIRP/LBP/DCI-PTP	70.13±0.96	29.70	600
[LIU et collab., 2016b]	MRELBP	73.43±0.84	32.52	800
[GUO et collab., 2010b]	LBPV	60.40±1.58	24.24	54
[NGUYEN et collab., 2016]	MC-SBP2	59.56±0.89	29.09	4000
[RATAJCZAK et collab., 2019]	LS-LCoLBP*	69.40*	—	90
[QI et collab., 2013]	MSJLBP	61.76±1.23	25.15	3540
[QI et collab., 2014]	PRICoLBP	64.80±1.43	24.68	4000
[LIU et collab., 2012b]	NI/RD/CI	49.06±0.85	19.91	2200
[LIU et collab., 2016a]	ELBP ₁₃₄₆	58.46±1.45	27.57	80
[LIU et collab., 2016a]	ELBP ₁₂₃₄₅₆	58.12±1.15	26.17	100
[LIU et collab., 2016a]	ELBP ₁₂₃₄₅₆	60.82±1.07	27.14	120
GUO et collab. [2010a]	CLBP-S/M/C	51.18±1.70	27.17	2200
ZHAO et collab. [2012b]	CLBC-S/M/C	56.93±1.04	27.15	1990
LIU et collab. [2014]	BRINT-CS-CM	62.60±1.10	28.57	360
[proposé]	circ-SRBP _{MinMaxMeanVar}	76.81±2.76	26.42	2048
[proposé]	circ-SRBP	76.90±2.83	26.49	2560
[proposé]	ring-SRBP _{MinMaxMeanVar}	74.53±0.68	31.90	4096
[proposé]	ring-SRBP	74.78±0.76	32.03	5120
[proposé]	circ-SMBP _{MinMaxMeanVar}	81.21±2.01	31.43	2048
[proposé]	circ-SMBP	81.38±2.23	31.50	2560
[proposé]	ring-SMBP _{MinMaxMeanVar}	73.43±0.99	33.78	256
[proposé]	ring-SMBP	74.70±0.86	33.94	320

Pour la base AFF, on reprends le cas de la similarité inter-espèces entre les images des classes #3, #6 qui présentent des structures d'écorce assez similaires et difficiles à distinguer même à l'œil nu. Cette ambiguïté de similarité était d'ailleurs rapportée dans [WENDEL et collab., 2011]. Afin de voir de plus près le comportement des descripteurs circ-SMBP et MRELBP [LIU et collab., 2016b], nous présentons les matrices de confusions relatives à ces descripteurs dans la Figure 6.13. On peut voir que le descripteur proposé circ-SMBP arrive à bien distinguer entre les écorces des espèces #3, #6 mieux que le descripteur MRELBP. Ceci est d'ailleurs confirmé par l'exemple d'une image requête de l'espèce #3 et les cinq premières images réponses les plus similaires selon circ-SMBP (c.f. Figure 6.14a) et selon MRELBP [LIU et collab., 2016b] (c.f. Figure 6.14b).

Au final, il est important de noter le cas de l'espèce #11 qui est la plus faiblement identifiée dans la base AFF. En effet, les images de l'espèce #11 ont pour la plupart la surface couverte par les lichens ce qui cache la structure de l'écorce (c.f. Figure 6.15). Cependant, pour le MRELBP [LIU et collab., 2016b] essentiellement basé sur le calcul des patchs médians, qui peuvent réduire l'effet du bruit et représenter ainsi une information pertinente moins sensible. À l'encontre de cela, notre descripteur circ-SMBP encode toute l'information des niveaux d'intensité, et dans ce cas celles des

6. Résultats expérimentaux 6.2. Évaluation des descripteurs de macro-motifs binaires statistiques

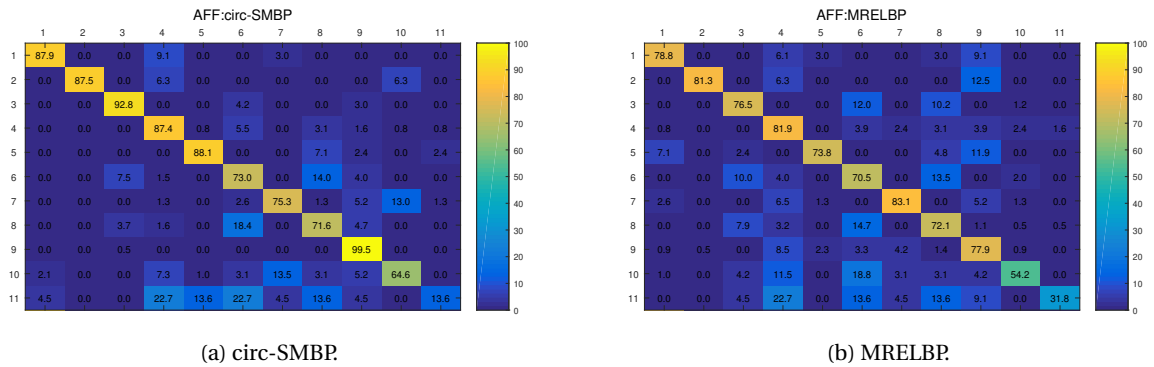


FIGURE 6.13 – Matrices de confusion des descripteurs circ-SMBP et MRELBP [Liu et collab., 2016b] sur la base AFF

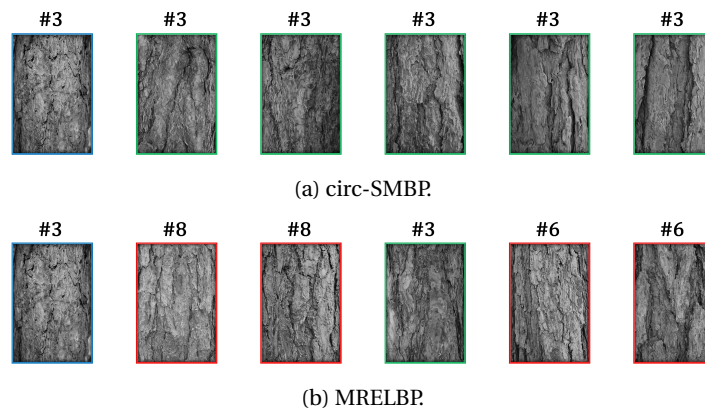


FIGURE 6.14 – Exemple de la même image requête de l'espèce #3 de la base AFF et ses réponses selon circ-SMBP et MRELBP.

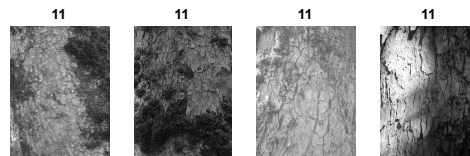


FIGURE 6.15 – Exemple d'écorce de l'espèce #11 de la base AFF dont la surface est couverte de lichen.

lichens aussi, lors de la représentation de la texture ce qui dégrade ses performances.

Toutefois, globalement, nos descripteurs sont assez efficaces pour l'identification des différentes espèces végétales dans des conditions naturelles plus ou moins modérées, et sont remarquablement discriminants contre le problème de similarité inter-espèces.

6.2.3 Résultats sur la base Bark101

6.2.3.1 Évaluation des paramètres

Pour l'évaluation de la représentation de la macrostructure sur la base Bark101, nous présentons dans les Tableaux 6.17 et 6.18 les scores de classification en 1-NN en fonction des différents niveaux d'échelle \mathcal{L} et des distances radiales τ . De façon similaire aux bases Trunk12 et AFE, on trouve que les descripteurs SMBP obtiennent les meilleures performances par rapport aux descripteurs radiaux. Cependant ces résultats restent très faibles pour les différentes représentations, soit en *circ* ou en *ring* comme le montre les courbes de la Figure 6.16 et les évaluations pour les différentes largeurs d'anneaux dans le Tableau 6.19.

Les configurations optimales sont présentées dans le Tableau 6.20 pour un score maximum obtenu avec le descripteur circ-SMBP avec 33.66% en 1-NN, et 37.37% en SVM_{rbf}.

Nous passons ensuite à la représentation des descripteurs SRBP et SMBP en différentes combinaisons à base de deux descripteurs (c.f. Figure 6.17) et les combinaisons de trois et quatre descripteurs (c.f. Tableau 6.21).

On trouve que c'est la combinaison des cinq descripteurs qui obtient les scores de classification les plus élevés. Cependant, il y a une stabilité dans le comportement de nos descripteurs en terme de complémentarité des descripteurs statistiques à représenter la macrostructure des différentes espèces d'écorce. Le classifieur SVM_{rbf} permet d'améliorer d'avantage les scores de classification par rapport au classifieur non paramétrique 1-NN.

6.2.3.2 Comparaison avec l'état de l'art

Une extension de l'étude expérimentale de la classification des espèces végétales à partir des images d'écorce est exposée dans cette section sur la base Bark101 avec plus de 100 espèces différentes d'écorce. Le Tableau 6.22 montre les résultats comparatifs obtenus avec les classifieurs 1-NN et SVM_{rbf}.

On peut voir qu'avec les deux classifieurs, tous nos descripteurs obtiennent les meilleurs scores sur la base Bark101, notamment pour circ-SMBP_{MinMaxMeanVar} avec 34.05% en 1-NN, et 36.60% en SVM_{rbf}. Cependant, malgré que les résultats de classification obtenus par nos descripteurs dépassent ceux obtenus par des méthodes en comparaison, ils restent considérablement faibles (moins de 40%) pour une application de classification d'un contenu visuel. Ceci peut être expliqué par le fait que pour la base Bark101, on a un très grand nombre de classes d'espèces (101) contre un faible nombre d'images d'écorce par espèce (2 – 138), avec une variabilité intra-espèce très importante. Ceci peut engendrer des confusions entre les images d'écorce des différentes espèces.

Par ailleurs, on remarque une consistance du comportement de notre approche sur des bases d'écorce à grande échelle. En effet, sur Bark101, on trouve que nos descripteurs en profondeur retournent des scores plus intéressants que ceux retournés par les descripteurs radiaux.

TABLEAU 6.17 – Taux de classification (%) en 1-NN du descripteur circ-SRBP sur la base Bark101 avec différentes représentations de la macrostructure.

circ-SRBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	2	4	16	32	64	128	256	512	1024	2048	2	4	16	32	64	128	256	512	2	4	16	32	64	128
$R_1 = 1$	16.21	19.61	22.39	24.32	25.32	26.33	27.02	27.10	27.95	27.41	7.64	21.46	24.32	26.56	28.49	28.88	29.18	29.50	18.06	22.16	26.48	28.88	29.96	30.42
$R_1 = 2$	17.29	21.31	23.08	23.70	24.47	25.71	27.18	28.03	28.10	27.95	16.60	21.54	24.24	25.86	27.87	28.18	28.41	28.88	18.37	23.32	27.02	27.87	28.41	29.49
$R_1 = 3$	17.06	21.93	23.32	24.63	25.09	27.10	27.64	28.41	28.50	27.72	16.83	21.15	24.24	26.25	26.79	28.10	28.18	27.79	16.60	22.77	25.86	27.25	27.49	28.33
$R_1 = 4$	17.37	21.31	23.55	24.78	26.33	27.41	27.64	28.18	27.95	27.33	17.14	21.62	25.71	26.71	27.56	28.33	28.33	28.10	19.15	23.08	25.55	26.64	27.87	28.80
$R_1 = 5$	17.83	21.85	23.86	25.94	27.49	27.41	27.95	27.79	28.26	28.33	19.07	23.24	26.25	27.41	28.64	29.11	28.49	28.64	19.45	24.09	26.87	28.10	28.64	28.41
$R_1 = 6$	18.30	22.70	25.01	26.56	27.18	28.03	27.87	28.03	28.49	28.26	19.61	23.62	26.71	27.33	28.50	28.80	28.80	28.64	19.61	26.40	27.41	28.33	27.87	29.26
circ-SRBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)	2	4	16	32	64	128	256	512	1024	2048	2	4	16	32	64	128	256	512	2	4	16	32	64	128
$R_1 = 1$	16.60	19.22	21.00	22.54	23.78	24.32	25.09	25.79	25.71	25.25	16.91	20.54	22.93	24.16	25.25	26.40	26.87	26.79	17.22	20.54	23.93	25.71	27.10	27.56
$R_1 = 2$	16.98	20.38	21.85	23.39	24.24	24.24	25.25	26.02	25.71	25.71	17.06	20.23	23.08	23.78	24.40	24.47	25.09	25.63	17.14	21.46	24.09	25.17	25.79	26.17
$R_1 = 3$	16.98	20.69	23.08	24.40	24.24	25.17	26.02	26.25	25.86	25.25	17.14	21.85	23.16	24.01	25.01	26.10	25.94	26.64	16.13	21.23	23.55	24.09	24.24	24.47
$R_1 = 4$	16.91	19.76	22.47	23.16	24.71	24.94	25.79	25.94	25.71	25.17	16.21	21.00	22.70	24.09	25.17	25.55	26.79	27.10	16.75	20.54	23.86	24.32	24.86	25.63
$R_1 = 5$	16.67	19.92	22.54	23.39	24.24	24.47	24.63	24.86	25.01	25.09	17.99	21.31	23.24	24.09	25.86	26.56	27.02	26.40	16.91	21.15	23.24	24.78	25.55	25.63
$R_1 = 6$	17.60	20.23	22.54	23.62	24.24	24.16	24.63	25.01	24.94	24.47	17.76	22.00	24.16	25.25	26.56	26.87	26.94	27.18	17.68	21.46	22.77	24.63	25.09	25.48

TABLEAU 6.18 – Taux de classification (%) en 1-NN du descripteur circ-SMBP sur la base Bark101 avec différentes représentations de la macrostructure.

circ-SMBP																									
<i>AllN</i>		$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}		2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)		2	4	16	32	64	128	256	512	1024	2048	2	4	16	32	64	128	256	512	2	4	16	32	64	128
$R_1 = 1$		14.13	21.00	24.71	27.02	28.49	29.03	29.42	29.42	29.88	29.42	6.71	24.40	27.18	28.88	30.19	31.73	32.12	32.58	13.12	24.63	28.95	31.35	32.12	32.58
$R_1 = 2$		13.43	23.24	25.25	26.56	28.10	29.18	29.88	29.57	29.80	29.65	12.97	25.40	28.49	29.72	31.04	31.81	31.96	32.81	11.81	26.40	30.11	31.89	33.05	33.05
$R_1 = 3$		13.20	24.55	25.63	27.64	28.49	28.95	30.27	30.19	30.27	29.80	13.28	25.17	28.33	30.11	31.58	31.89	32.66	32.81	12.50	27.33	30.81	31.27	32.81	33.51
$R_1 = 4$		15.21	23.93	26.40	27.72	29.26	29.18	30.73	30.65	30.73	30.73	13.43	26.40	28.57	30.81	31.66	32.66	32.89	33.28	13.20	27.18	30.42	32.04	33.51	33.28
$R_1 = 5$		14.28	24.01	26.71	28.64	28.88	29.65	30.88	31.66	31.27	31.35	12.50	25.86	28.72	31.11	32.12	32.43	33.12	33.66	12.35	26.79	31.73	33.28	33.51	32.81
$R_1 = 6$		14.36	24.32	27.87	28.95	29.18	30.34	30.73	31.58	31.90	31.73	13.12	25.79	29.49	32.12	32.43	33.28	33.43	33.43	12.12	28.57	32.50	32.43	32.66	32.50
circ-SMBP																									
<i>OneN</i>		$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}		2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim ($2^{\mathcal{L}} \times 5$)		2	4	16	32	64	128	256	512	1024	2048	2	4	16	32	64	128	256	512	2	4	16	32	64	128
$R_1 = 1$		13.43	19.92	22.70	24.47	25.86	26.17	26.56	27.10	26.87	27.56	13.66	20.84	24.55	26.25	27.41	28.26	28.57	28.80	14.59	22.23	26.48	27.87	28.57	29.65
$R_1 = 2$		15.98	21.62	23.47	26.02	26.87	26.64	26.64	27.18	27.49	27.56	15.44	23.16	25.09	27.25	28.03	28.49	28.57	29.11	13.43	23.16	27.72	28.41	29.96	29.80
$R_1 = 3$		13.74	21.00	24.09	25.86	26.48	27.18	27.41	27.64	27.72	27.41	12.43	23.16	26.40	27.25	28.41	28.95	29.03	29.11	12.43	23.08	26.71	28.72	29.72	29.49
$R_1 = 4$		15.05	21.69	23.32	25.71	25.94	27.41	27.56	27.02	27.87	27.49	12.97	21.39	24.63	26.10	27.95	27.95	28.33	28.26	13.89	24.16	27.10	28.18	29.03	28.88
$R_1 = 5$		14.05	20.46	23.55	25.09	26.10	26.64	27.18	28.03	28.10	27.87	16.98	23.93	26.40	27.64	28.26	28.18	27.95	28.10	13.66	24.09	27.33	27.79	28.64	29.18
$R_1 = 6$		14.59	20.92	24.24	24.40	26.02	26.94	27.72	27.95	28.18	27.64	13.66	24.09	27.10	27.64	27.64	27.64	27.79	28.41	13.59	24.47	27.02	28.26	28.49	28.41

6. Résultats expérimentaux 6.2. Évaluation des descripteurs de macro-motifs binaires statistiques

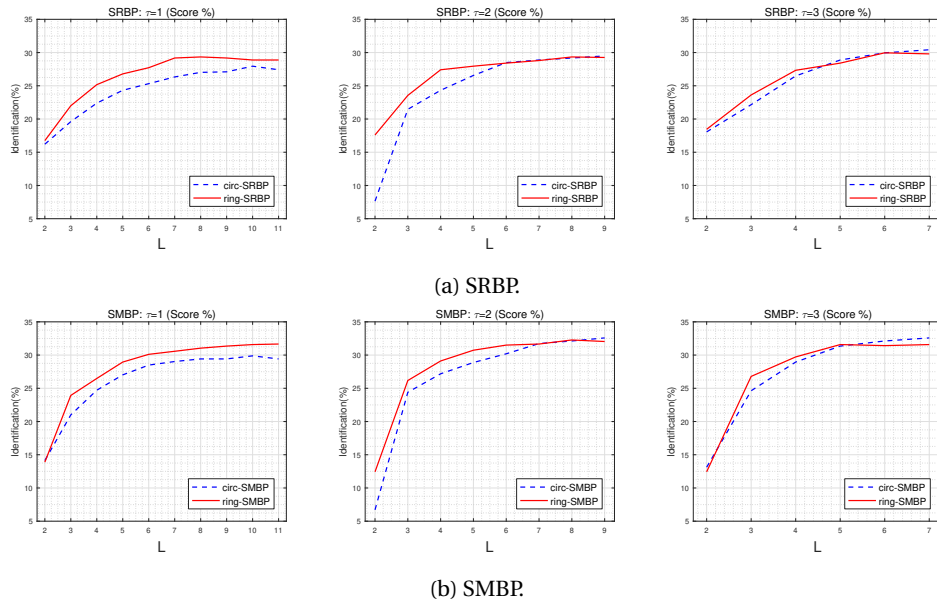


FIGURE 6.16 – Comparaison des scores d'identification en 1-NN et SVM_{rbf} pour les descripteurs SRBP et SMBP avec les configurations *circ* et *ring* sur la base Bark101.

TABLEAU 6.19 – Taux de classification (%) en 1-NN des descripteurs ring-SRBP et ring-SMBP sur la base Bark101 avec différentes représentations de la macrostructure.

ring-SRBP										
<i>AllN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	16.75	22.00	25.17	26.79	27.72	29.18	29.34	29.18	28.88	28.88
$\tau = 3$	17.60	23.55	27.41	27.95	28.41	28.80	29.34	29.26	28.41	27.79
$\tau = 4$	18.45	23.62	27.33	28.41	29.96	29.80	30.27	29.96	29.34	28.49
$\tau = 5$	19.84	24.40	26.40	28.95	28.57	28.10	28.18	27.41	27.10	25.55
<i>OneN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	17.06	20.07	22.62	24.86	25.79	26.40	27.18	26.33	27.18	26.40
$\tau = 3$	18.14	22.47	25.71	27.02	27.41	27.49	27.87	27.95	27.18	26.71
$\tau = 4$	18.22	23.62	26.79	26.94	26.87	27.10	28.33	28.49	27.10	26.79
$\tau = 5$	19.15	24.63	26.17	27.41	28.49	28.57	27.49	28.03	27.95	26.71
ring-SMBP										
<i>AllN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	13.89	23.93	26.48	28.95	30.11	30.57	31.04	31.35	31.58	31.66
$\tau = 3$	12.43	26.17	29.11	30.73	31.50	31.66	32.27	32.04	31.11	30.96
$\tau = 4$	12.43	26.79	29.72	31.58	31.42	31.58	31.50	29.80	28.72	28.03
$\tau = 5$	11.73	27.10	29.72	31.27	31.73	28.64	28.03	28.18	27.49	27.41
<i>OneN</i>										
\mathcal{L}	2	3	4	5	6	7	8	9	10	11
Dim ($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048
$\tau = 2$	15.13	23.47	25.25	26.87	28.03	28.88	29.96	29.80	29.49	29.03
$\tau = 3$	14.44	23.70	28.03	29.11	31.42	32.43	31.58	31.11	30.81	30.03
$\tau = 4$	13.66	25.01	29.72	30.96	31.04	30.88	31.19	29.34	28.41	27.33
$\tau = 5$	14.36	27.33	30.11	29.80	30.50	28.49	27.18	27.49	27.56	26.48

TABLEAU 6.20 – Configurations retenues pour les différents descripteurs statistiques sur la base Bark101.

Descripteur	Configuration	1-NN	SVM_{rbf}	Dim	Descripteur	Configuration	1-NN	SVM_{rbf}	Dim
circ-SRBP	$\{\mathcal{L} = 7, R_1 = 1, \tau = 3, AllN\}$	30.42	36.83	640	ring-SRBP	$\{\mathcal{L} = 8, \tau = 4, AllN\}$	30.27	37.29	1280
circ-SRBP	$\{\mathcal{L} = 7, R_1 = 1, \tau = 3, OneN\}$	27.56	35.29	640	ring-SRBP	$\{\mathcal{L} = 7, \tau = 5, OneN\}$	28.57	35.90	640
Descripteur	Configuration	1-NN	SVM_{rbf}	Dim	Descripteur	Configuration	1-NN	SVM_{rbf}	Dim
circ-SMBP	$\{\mathcal{L} = 9, R_1 = 5, \tau = 2, AllN\}$	33.66	37.29	2560	ring-SMBP	$\{\mathcal{L} = 8, \tau = 3, AllN\}$	32.27	36.44	1280
circ-SMBP	$\{\mathcal{L} = 6, R_1 = 2, \tau = 3, OneN\}$	29.96	37.06	320	ring-SMBP	$\{\mathcal{L} = 7, \tau = 3, OneN\}$	32.43	36.52	640

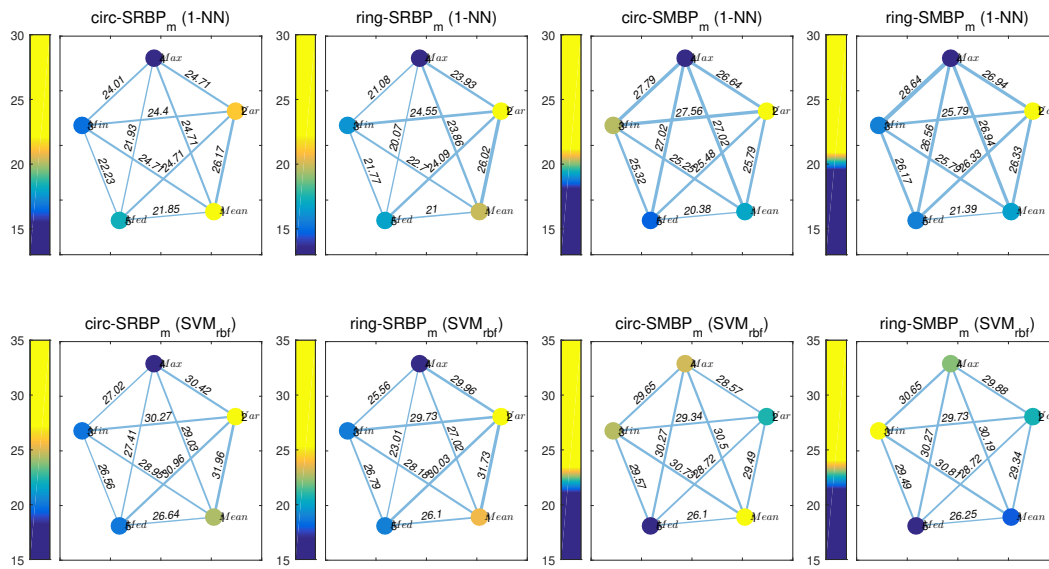


FIGURE 6.17 – Taux de classification (%) en 1-NN et SVM_{rbf} des combinaisons formées de deux descripteurs statistiques sur la base Bark101.

TABLEAU 6.21 – Taux de classification (%) en 1-NN et SVM_{rbf} des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base Bark101.

circ-SRBP ($\mathcal{L} = 7$)				ring-SRBP ($\mathcal{L} = 8$)			
Descripteur	1-NN	SVM_{rbf}	Dim	Descripteur	1-NN	SVM_{rbf}	Dim
SRBP _{MaxMeanVar}	28.03	35.21	384	SRBP _{MinMedVar}	26.79	33.28	768
SRBP _{MinMaxMean}	28.33	32.74	384	SRBP _{MinMeanVar}	27.18	35.52	768
SRBP _{MinMeanVar}	28.95	34.05	384	SRBP _{MaxMeanVar}	27.79	34.20	768
SRBP _{MinMaxMeanMed}	28.80	33.59	512	SRBP _{MinMeanMedVar}	28.18	36.37	1024
SRBP _{MinMeanMedVar}	29.18	35.29	512	SRBP _{MinMaxMedVar}	28.95	35.13	1024
SRBP _{MinMaxMeanVar}	30.34	36.83	512	SRBP _{MinMaxMeanVar}	29.34	36.91	1024
SRBP	30.42	36.83	640	SRBP	30.27	37.37	1280
circ-SMBP ($\mathcal{L} = 9$)				ring-SMBP ($\mathcal{L} = 8$)			
Descripteur	1-NN	SVM_{rbf}	Dim	Descripteur	1-NN	SVM_{rbf}	Dim
SMBP _{MinMaxVar}	31.35	34.44	1536	SMBP _{MinMedVar}	30.11	33.05	768
SMBP _{MinMaxMean}	32.20	34.98	1536	SMBP _{MinMaxMed}	30.42	33.05	768
SMBP _{MinMaxMed}	32.20	32.97	1536	SMBP _{MinMaxMean}	31.66	33.66	768
SMBP _{MinMaxMeanMed}	31.74	35.05	2048	SMBP _{MinMaxMeanVar}	31.35	35.90	1024
SMBP _{MinMaxMedVar}	32.81	34.74	2048	SMBP _{MinMaxMedVar}	31.58	34.74	1024
SMBP _{MinMaxMeanVar}	34.05	36.60	2048	SMBP _{MinMaxMeanMed}	31.73	35.52	1024
SMBP	33.66	37.29	2560	SMBP	32.27	36.44	1280

TABLEAU 6.22 – Résultats de comparaison avec l'état de l'art sur la base Bark101. Les résultats marqués par * sont pris de [RATAJCZAK et collab., 2019].

Ref	Descripteur	1-NN	SVM _{rbf}	Dim
[WAN et collab., 2004]	GLCM	05.94	05.25	20
[HUANG et collab., 2006d]	Gabor	22.70	05.32	48
[HUANG et collab., 2006b]	DWT	13.89	14.51	19
[OJALA et collab., 2002]	LBP	15.90	16.37	256
[OJALA et collab., 2002]	MResLBP	21.46	30.81	54
[ŠULC et MATAS, 2013]	MSLBP	25.32	30.76	180
[RATAJCZAK et collab., 2019]	LS-LCoLBP*	28.30*	30.10*	90
[proposé]	circ-SRBP _{MinMaxMeanVar}	30.34	36.83	512
[proposé]	circ-SRBP	30.42	36.83	640
[proposé]	ring-SRBP _{MinMaxMeanVar}	29.34	36.91	1024
[proposé]	ring-SRBP	30.27	37.37	1280
[proposé]	circ-SMBP _{MinMaxMeanVar}	34.05	36.60	2048
[proposé]	circ-SMBP	33.66	37.29	2560
[proposé]	ring-SMBP _{MinMaxMeanMed}	31.73	35.52	1024
[proposé]	ring-SMBP	32.27	36.44	1280

6.2.4 Résultats sur la base BarkNet

6.2.4.1 Évaluation des paramètres

Une extension de l'étude expérimentale de la classification des espèces végétales à partir des images d'écorce est exposée dans cette section sur une très grande base d'écorce. Il s'agit de la base BarkNet avec plus de 23000 images d'écorce.

Les Tableaux 6.23 et 6.24 présentent l'évaluation des scores de classification des espèces de la base BarkNet en fonction des niveaux d'échelle \mathcal{L} et des distances radiales τ pour une macro-structure *circ*, et le Tableau 6.25 présente les scores d'évaluation pour une macrostructure *ring*. Ces résultats expriment les taux de classification en 1-NN.

On peut voir à partir de ces tableaux, que c'est la représentation circulaire *circ* de la macro-structure qui obtient les meilleurs scores de classification par rapport à une représentation en anneaux *ring*. De plus, ces résultats sont obtenus avec des valeurs de distance radiale (respectivement largeur d'anneau) τ faible, à l'encontre des autres bases d'écorce où on trouve que c'est les valeurs élevées de τ qui retournent des scores élevés.

Cela revient essentiellement à la taille des images de la base BarkNet. En effet les images de BarkNet sont d'une résolution très faible par rapport à celles des autres bases d'écorce, i.e., 224×224 contre des tailles > 1000 pixels selon au moins une dimension de l'image. De ce fait, plus la taille du support spatial de la macrostructure est grande pour une image de faible résolution, moins il y a de pixels à encoder et donc pas suffisamment d'information de texture pour caractériser le macro-motif relatif de l'image. Ceci est d'ailleurs confirmé par les courbes de la Figure 6.18 qui représentent les scores d'une macrostructure *circ* pour les différentes valeurs τ et celles correspondantes en largeur d'anneau pour les macrostructures *ring*. Plus la valeur de l'échelle \mathcal{L} augmente plus les scores de la représentation en *ring* diminuent, de même pour les grandes valeurs de τ .

Les configurations optimales sur la base BarkNet sont présentées dans le Tableau 6.26. Nous remarquons un comportement stable de nos descripteurs quant à la représentation par les descripteurs profonds par rapport aux descripteurs radiaux comme trouvé sur les autres bases d'écorce.

On passe ensuite aux combinaisons des descripteurs statistiques pour les deux encodages : radial et profond, en utilisant des combinaisons de deux, trois et quatre descripteurs. La Figure 6.19 montre les scores de classification avec 1-NN et SVM_{lin} des combinaisons à base de deux descripteurs, et le Tableau 6.27 présente ceux des combinaisons de plus de trois descripteurs. Une fois de plus, nous trouvons que plus on combine des descripteurs statistiques, plus on améliore les performances de la classification des espèces d'écorce, d'où la complémentarité des descriptions statistiques. Aussi, c'est les descripteurs qui contiennent les approximations des moments du premier et second ordre qui obtiennent les meilleurs scores.

TABLEAU 6.23 – Taux de classification (%) en 1-NN du descripteur circ-SRBP sur la base BarkNet avec différentes représentations de la macrostructure.

circ-SRBP																								
<i>AllN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	36.01	38.95	42.58	46.16	48.24	48.12	48.35	49.04	49.13	48.06	35.10	42.10	45.00	46.19	46.21	45.48	45.32	43.15	36.06	43.45	43.59	44.20	43.02	42.47
$R_1 = 2$	34.35	39.48	45.00	46.82	47.14	48.90	48.38	47.99	47.81	47.24	34.69	42.76	44.20	45.57	45.20	45.71	44.47	42.60	35.12	41.33	43.11	43.47	42.92	41.01
$R_1 = 3$	34.62	42.06	45.87	47.03	47.81	48.15	48.01	47.76	47.17	46.57	34.30	41.60	44.36	44.25	44.06	43.79	42.63	42.35	34.62	41.44	43.20	43.27	42.10	40.21
$R_1 = 4$	36.29	41.97	45.64	46.94	48.08	47.92	47.76	48.08	47.76	45.46	35.90	40.87	43.49	43.95	44.45	43.75	43.08	41.05	35.33	40.21	41.58	40.96	41.60	40.28
$R_1 = 5$	36.47	42.08	45.39	46.25	46.82	47.53	47.37	47.35	46.57	44.84	35.67	41.30	43.56	44.02	43.75	42.54	41.26	39.53	34.83	39.96	41.85	40.92	40.23	39.23
$R_1 = 6$	34.69	41.35	44.95	46.05	46.85	46.09	46.23	46.35	44.98	43.77	34.42	41.74	42.90	42.83	42.35	41.74	40.64	40.46	33.73	40.44	40.71	40.57	39.37	39.09

circ-SRBP																								
<i>OneN</i>	$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}	2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim($2^{\mathcal{L}} \times 5$)	4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$	36.24	40.48	43.70	45.11	47.44	48.56	49.27	48.90	49.11	48.60	34.76	40.92	45.32	47.39	46.46	47.67	47.03	46.55	36.13	41.78	43.36	45.57	44.25	43.38
$R_1 = 2$	33.57	39.46	43.04	47.19	48.47	49.20	49.50	49.72	49.38	49.56	34.46	41.71	45.36	46.12	46.55	46.57	46.55	45.32	35.35	41.81	44.82	44.77	43.86	43.77
$R_1 = 3$	33.96	41.65	46.57	48.47	49.52	50.06	49.77	50.18	49.31	48.81	34.19	43.95	46.53	46.30	46.37	46.82	46.76	45.55	34.83	40.73	43.08	42.92	42.72	42.88
$R_1 = 4$	36.01	43.04	47.01	48.31	49.01	49.70	49.77	49.29	48.47	47.96	33.85	41.30	43.72	45.71	46.60	45.71	45.71	44.86	34.55	38.27	41.51	41.76	41.56	40.69
$R_1 = 5$	35.05	41.65	44.93	46.25	47.51	48.12	48.76	48.01	47.17	46.00	35.88	42.22	43.81	44.54	44.75	44.45	44.54	43.54	34.30	39.00	40.16	41.58	42.06	40.76
$R_1 = 6$	34.55	41.76	44.68	46.41	46.35	46.92	47.39	45.80	46.32	44.66	33.85	40.64	43.43	45.07	44.34	44.16	42.99	42.08	31.82	37.56	38.98	39.96	40.30	39.55

TABLEAU 6.24 – Taux de classification (%) en 1-NN du descripteur circ-SMBP sur la base BarkNet avec différentes représentations de la macrostructure.

circ-SMBP																									
<i>AllN</i>		$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}		2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim($2^{\mathcal{L}} \times 5$)		4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$		27.03	40.60	45.89	47.78	50.27	51.32	51.11	51.18	50.68	50.72	24.52	43.88	48.01	48.44	48.54	47.74	47.83	46.62	23.69	43.75	45.66	45.96	46.55	45.55
$R_1 = 2$		26.77	40.57	47.14	49.63	51.41	50.75	51.25	50.77	50.06	49.47	23.90	44.18	47.24	47.78	46.96	46.80	46.60	46.53	24.01	42.33	44.34	44.98	44.75	43.54
$R_1 = 3$		24.01	42.19	48.01	49.77	50.45	49.86	50.38	49.22	49.58	48.12	23.38	42.70	46.44	46.48	45.87	45.59	45.18	45.30	22.42	41.46	43.61	44.06	43.72	41.78
$R_1 = 4$		25.31	43.36	48.54	48.81	47.99	48.69	48.35	48.08	46.98	47.19	23.97	41.97	46.78	46.55	45.48	44.77	44.45	42.90	23.26	40.00	43.02	42.51	42.06	40.28
$R_1 = 5$		23.65	41.26	46.03	47.21	47.55	47.35	47.60	46.87	46.85	47.03	22.49	41.42	43.59	43.45	43.95	43.63	42.29	41.37	22.74	38.41	41.21	41.46	40.19	38.80
$R_1 = 6$		23.17	41.90	45.41	46.98	46.41	47.35	46.69	46.69	46.82	46.23	22.05	40.00	42.26	43.54	42.47	42.86	42.33	40.94	21.32	38.61	40.85	40.30	39.07	38.68
circ-SMBP																									
<i>OneN</i>		$\tau = 1$										$\tau = 2$									$\tau = 3$				
\mathcal{L}		2	3	4	5	6	7	8	9	10	11	2	3	4	5	6	7	8	9	2	3	4	5	6	7
Dim($2^{\mathcal{L}} \times 5$)		4	8	16	32	64	128	256	512	1024	2048	4	8	16	32	64	128	256	512	4	8	16	32	64	128
$R_1 = 1$		27.30	38.43	44.52	46.66	48.85	49.68	51.02	51.14	51.07	50.02	26.16	41.78	45.73	47.01	48.03	47.71	47.94	46.94	25.91	42.47	44.11	44.66	44.93	44.50
$R_1 = 2$		27.82	39.78	46.30	48.10	50.31	51.07	51.09	50.84	50.20	49.88	28.26	42.40	45.16	46.62	47.67	46.64	46.37	46.16	25.18	41.90	43.52	44.43	44.59	44.13
$R_1 = 3$		25.75	39.64	46.09	48.92	49.93	49.95	50.66	50.82	49.38	48.49	24.24	42.88	46.85	46.94	46.78	46.82	46.39	45.96	24.49	41.05	43.59	42.72	43.15	42.10
$R_1 = 4$		26.23	41.42	46.82	47.83	49.84	49.81	50.27	49.01	48.40	48.54	24.47	38.75	43.90	45.16	45.71	45.84	45.23	44.32	24.40	39.71	40.80	42.19	43.06	41.37
$R_1 = 5$		26.57	39.32	44.91	46.14	47.76	47.58	46.57	46.80	47.51	47.08	24.65	41.65	43.36	44.79	44.52	45.23	45.30	44.89	24.40	36.92	40.30	41.60	42.19	41.28
$R_1 = 6$		26.18	37.15	42.54	45.73	45.20	46.32	46.85	46.64	47.60	45.82	24.77	39.96	43.24	43.52	44.41	43.86	43.77	43.49	23.54	34.58	34.58	40.89	42.08	40.87

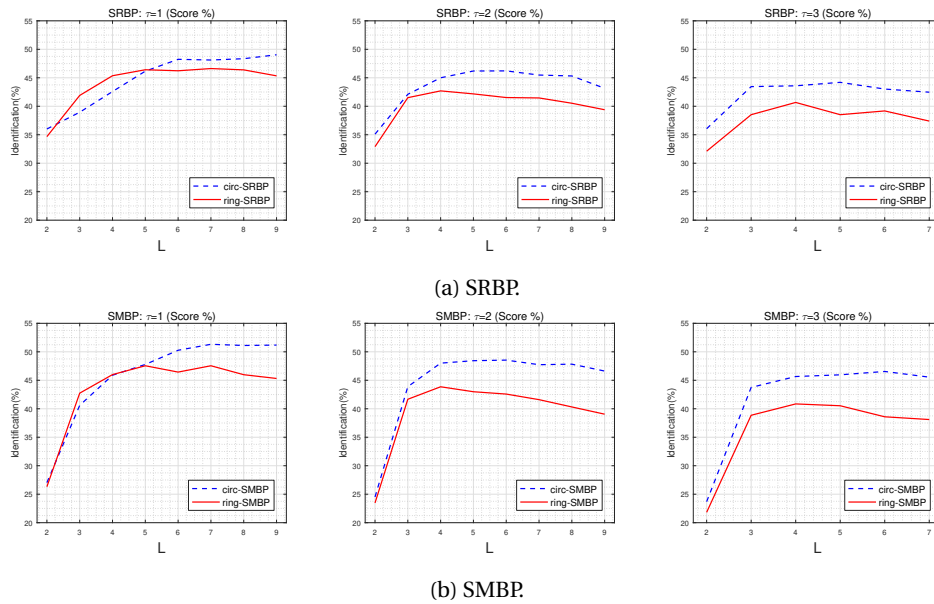


FIGURE 6.18 – Comparaison des scores de classification (%) en 1-NN et SVM_{lin} pour les descripteurs SRBP et SMBP avec les configurations *circ* et *ring* sur la base BarkNet.

6.2.4.2 Comparaison avec l'état de l'art

Le Tableau 6.28 montre les résultats comparatifs obtenus sur la base BarkNet avec les classifieurs 1-NN et SVM_{lin} . Le descripteur proposé circ-SRBP retourne 50.18% en 1-NN et 57.27% en SVM_{lin} contre MSLBP [ŠULC et MATAS, 2013] avec 44.75% 48.63% avec les mêmes classifieurs. Pour un aperçu visuel, nous avons projeté la matrice de confusion de notre descripteur circ-SRBP et celle du descripteur MSLBP [ŠULC et MATAS, 2013] sur la base BarkNet, comme le montre la Figure 6.20. On peut voir que sur l'ensemble des images de la base BarkNet, le circ-SRBP réduit la confusion entre les différentes espèces. Toutefois, cela reste insuffisant en raison du très grand nombre d'images que contient cette base pour 20 classes d'espèces (23000 images), et une similarité inter-espèces notable due à la présence de plusieurs espèces d'écorce appartenant à la même famille d'espèces [CARPENTIER et collab., 2018].

Par ailleurs, cette expérience nous a permis non seulement d'évaluer l'extension de nos descripteurs sur des bases assez défiantes, mais aussi de mesurer leurs limites quand à la reconnaissance des espèces végétales.

TABLEAU 6.25 – Taux de classification (%) en 1-NN des descripteurs ring-SRBP et ring-SMBP sur la base BarkNet avec différentes représentations de la macrostructure.

ring-SRBP									ring-SMBP								
<i>AllN</i>									<i>OneN</i>								
\mathcal{L}	2	3	4	5	6	7	8	9	\mathcal{L}	2	3	4	5	6	7	8	9
Dim (\mathcal{L})	4	8	16	32	64	128	256	512	Dim (\mathcal{L})	4	8	16	32	64	128	256	512
$\tau = 2$	34.67	41.90	45.36	46.41	46.23	46.62	46.39	45.34	$\tau = 2$	33.48	42.22	45.34	46.60	46.94	47.21	47.10	46.41
$\tau = 3$	32.91	41.51	42.70	42.17	41.53	41.46	40.53	39.39	$\tau = 3$	32.89	42.31	44.86	45.73	44.91	44.75	43.13	41.71
$\tau = 4$	32.11	38.52	40.67	38.52	39.18	37.40	36.70	35.37	$\tau = 4$	33.89	40.99	42.74	42.51	41.44	39.64	37.97	35.78
$\tau = 5$	33.71	37.75	38.70	38.29	35.81	34.71	33.34	30.83	$\tau = 5$	34.28	39.75	39.09	40.07	37.97	36.90	33.73	30.45

TABLEAU 6.26 – Configurations retenues pour les différents descripteurs statistiques sur la base BarkNet.

Descripteur	Configuration	1-NN	SVM _{lin}	Dim	Descripteur	Configuration	1-NN	SVM _{lin}	Dim
circ-SRBP	{ $\mathcal{L} = 10, R_1 = 1, \tau = 1, AllN$ }	49.13	57.09	5120	ring-SRBP	{ $\mathcal{L} = 7, \tau = 2, AllN$ }	46.62	57.61	640
circ-SRBP	{ $\mathcal{L} = 9, R_1 = 3, \tau = 1, OneN$ }	50.18	57.27	2560	ring-SRBP	{ $\mathcal{L} = 7, \tau = 2, OneN$ }	47.21	57.43	640
Descripteur	Configuration	1-NN	SVM _{lin}	Dim	Descripteur	Configuration	1-NN	SVM _{lin}	Dim
circ-SMBP	{ $\mathcal{L} = 6, R_1 = 2, \tau = 1, AllN$ }	51.41	56.13	320	ring-SMBP	{ $\mathcal{L} = 7, \tau = 2, AllN$ }	47.55	54.90	640
circ-SMBP	{ $\mathcal{L} = 9, R_1 = 1, \tau = 1, OneN$ }	51.14	56.07	2560	ring-SMBP	{ $\mathcal{L} = 7, \tau = 2, OneN$ }	47.71	56.15	640

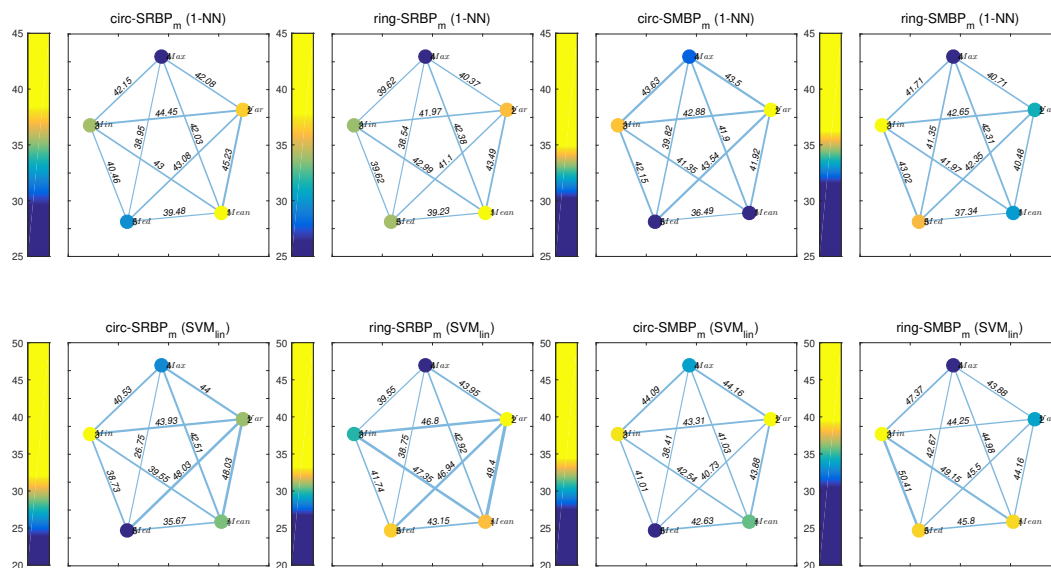


FIGURE 6.19 – Taux de classification (%) en 1-NN et SVM_{lin} des combinaisons formées de deux descripteurs statistiques sur la base BarkNet.

6. Résultats expérimentaux 6.2. Évaluation des descripteurs de macro-motifs binaires statistiques

TABLEAU 6.27 – Taux de classification (%) en 1-NN et SVM_{lin} des combinaisons formées de trois, quatre et tous les descripteurs statistiques sur la base BarkNet.

circ-SRBP ($\mathcal{L} = 9$)				ring-SRBP ($\mathcal{L} = 7$)			
Descripteur	1-NN	SVM _{lin}	Dim	Descripteur	1-NN	SVM _{lin}	Dim
SRBP _{MinMeanVar}	48.42	50.68	1536	SRBP _{MeanMedVar}	40.87	51.64	384
SRBP _{MinMaxVar}	46.09	51.52	1536	SRBP _{MaxMeanVar}	44.22	52.73	384
SRBP _{MaxMeanVar}	46.87	53.10	1536	SRBP _{MinMeanVar}	44.70	52.92	384
SRBP _{MinMaxMeanMed}	47.60	53.12	2048	SRBP _{MinMaxMedVar}	46.78	55.31	512
SRBP _{MinMeanMedVar}	48.51	53.65	2048	SRBP _{MinMeanMedVar}	44.38	55.70	512
SRBP _{MinMaxMeanVar}	49.54	55.13	2048	SRBP _{MinMaxMeanVar}	47.08	56.06	512
SRBP	50.18	57.27	2560	SRBP	46.62	57.61	640
circ-SMBP ($\mathcal{L} = 6$)				ring-SMBP ($\mathcal{L} = 7$)			
Descripteur	1-NN	SVM _{lin}	Dim	Descripteur	1-NN	SVM _{lin}	Dim
SMBP _{MinMeanMed}	44.73	48.97	192	SMBP _{MinMedVar}	46.16	52.37	384
SMBP _{MaxMeanVar}	47.62	49.88	192	SMBP _{MinMaxMean}	46.48	53.28	384
SMBP _{MinMaxVar}	48.33	50.45	192	SMBP _{MinMeanMed}	42.99	53.96	384
SMBP _{MinMaxMedVar}	51.11	52.69	256	SMBP _{MinMaxMeanMed}	47.35	54.35	512
SMBP _{MinMeanMedVar}	47.19	52.82	256	SMBP _{MinMaxMeanVar}	47.08	54.47	512
SMBP _{MinMaxMeanVar}	50.82	53.37	256	SMBP _{MinMeanMedVar}	45.05	55.08	512
SMBP	51.41	56.13	320	SMBP	47.71	56.15	640

TABLEAU 6.28 – Résultats de comparaison avec l'état de l'art sur la base BarkNet.

Ref	Descripteur	1-NN	SVM _{lin}	Dim
[WAN et collab., 2004]	GLCM	16.49	20.75	20
[HUANG et collab., 2006d]	Gabor	23.22	38.00	48
[HUANG et collab., 2006b]	DWT	28.76	40.55	19
[OJALA et collab., 2002]	LBP	30.26	32.02	256
[OJALA et collab., 2002]	MResLBP	42.70	47.33	54
[ŠULC et MATAS, 2013]	MSLBP	44.75	48.63	180
[proposé]	circ-SRBP _{MinMaxMeanVar}	49.54	55.13	2048
[proposé]	circ-SRBP	50.18	57.27	2560
[proposé]	ring-SRBP _{MinMaxMeanVar}	47.08	56.06	512
[proposé]	ring-SRBP	46.62	57.61	640
[proposé]	circ-SMBP _{MinMaxMeanVar}	50.82	53.37	256
[proposé]	circ-SMBP	51.41	56.13	320
[proposé]	ring-SMBP _{MinMeanMedVar}	45.05	55.08	512
[proposé]	ring-SMBP	47.71	56.15	640

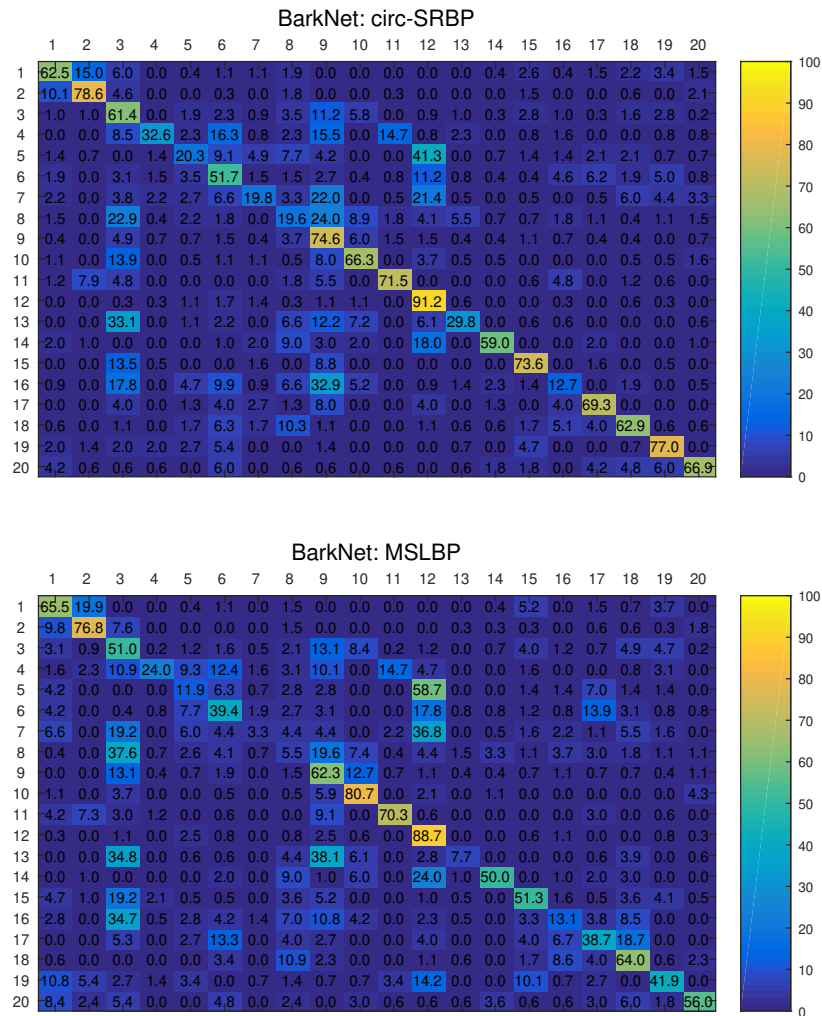


FIGURE 6.20 – Matrices de confusion des descripteurs circ-SMBP et MRELBP [LIU et collab., 2016b] avec SVM_{lin} sur la base BarkNet.

6.3 Classification des écorces par les CNN

En raison du succès remarquable des réseaux de neurones à convolution, CNN, dans diverses applications de classification d'images, nous avons tenu à introduire une évaluation de leurs performances ainsi qu'une étude comparative entre nos descripteurs statistiques de macro-motifs binaires, SRBP et SMBP.

Il est important de signaler que dans cette partie, il ne s'agit pas d'évaluer les performances de différents modèles CNN. Nous évaluons les performances du modèle ResNet34, comme dans le travail de [CARPENTIER et collab. \[2018\]](#) sur la base BarkNet. Aussi, Nous avons voulu étendre l'application de ResNet34 sur les trois autres bases, à savoir, Trunk12, AFF, et Bark101, et avoir un aperçu sur son comportement, bien que ces bases sont de tailles relativement petites par rapport à celle de BarkNet. Par la suite, nous établissons un comparatif entre nos vecteurs descripteurs à base d'attributs dits *handcrafted*, et les vecteurs descripteurs extraits à partir de ResNet34.

6.3.1 Détails d'implémentation

Les détails d'implémentation portent essentiellement sur : (i) le partage des bases d'images en ensemble d'entraînement et ensemble de test, (ii) les pré-traitements appliqués aux images, et (iii) les valeurs d'hyper-paramètres du modèle ResNet34.

Pour ce qui est du partage des bases, les ensembles d'entraînement et ceux de test étaient déjà définis pour les bases Bark101 et BarkNet par les auteurs originaux. Pour la base Bark101, 50% des images d'une classe sont aléatoirement choisies pour l'entraînement et le reste pour le test. Pour la base BarkNet, 80% des arbres pour l'entraînement et le 20% restant pour le test. Les trois classes ignorées ne contiennent pas suffisamment d'arbre [[CARPENTIER et collab., 2018](#)]. Pour les deux bases AFF et Trunk12, nous avons suivi le partage de Bark101 avec 50% des images d'une classe pour l'entraînement et le reste pour le test.

Nous avons conservé le pré-traitement précédent sur les images des bases Trunk12 et AFF, i.e., toutes les images sont redimensionnées à une largeur de 1000pixels de leurs tailles originales, en conservant le ratio. Pour les images de Bark101 et BarkNet, on garde les tailles initiales des images. Par ailleurs, pour toute image passée au ResNet34, on extrait aléatoirement un patch de 224×224 , étant la taille d'image en entrée au modèle. Une augmentation avec un retournement horizontal est appliquée aussi, comme réalisé par [CARPENTIER et collab. \[2018\]](#).

Au final, pour les valeurs des hyperparamètres, l'entraînement du CNN tourne initialement pour un total de 100 epoch avec un lot de taille 32, un taux d'apprentissage initial, $lr = 10^{-3}$. Ce taux est dégradé par un facteur de 5 à différentes epoch pour chacune des bases d'écorce testées. La méthode Adam est utilisée pour l'optimisation avec une dégradation des poids égale à 10^{-3} comme dans [[CARPENTIER et collab., 2018](#)].

Dans le cas d'utilisation d'un ResNet34 pré-entraîné sur la base *ImageNet*, nous avons conservé cette même configuration. De plus, la première couche du ResNet34 est gelée, i.e., les poids associés restent ceux appris sur la base *ImageNet*. Les couches supérieures sont réadaptées sur chacune des bases d'écorce, car les images sont très différentes de celles de la base *ImageNet*. Un taux d'apprentissage initial est fixé à $lr = 10^{-3}$. Ce taux est dégradé par un facteur de 5 aux epoch {16,33}. L'entraînement du réseau tourne pour un total de 40 epoch comme réalisé dans [[CARPEN-](#)

TABLEAU 6.29 – Résultats de la classification des espèces d'écorce avec ResNet34 entièrement entraîné sur les bases d'écorce.

Bases d'écorce	# Img	Epoch	Epoch de dégradation lr	Score (%)	Précision	Rappel	F1-score
Trunk12	393	22	15	49.74	0.62	0.49	0.54
AFF	1182	28	16	65.87	0.65	0.70	0.68
Bark101	2586	40	16, 33	18.61	0.15	0.19	0.17
BarkNet	23417	50	24, 36	73.41	0.71	0.74	0.73

TABLEAU 6.30 – Taux de classification des images d'écorce avec ResNet34 pré-entraîné et réadapté sur les bases d'écorce.

Bases d'écorce	# Img	Epoch	Epoch de dégradation lr	Score (%)	Précision	Rappel	F1-score
Trunk12	393	40	16, 33	90.76	0.90	0.90	0.90
AFF	1182	40	16, 33	89.13	0.91	0.92	0.92
Bark101	2586	40	16, 33	54.82	0.45	0.41	0.43
BarkNet	23417	40	16, 33	79.62	0.78	0.80	0.79

[TIER et collab., 2018](#)]. Cette configuration nous permet d'établir une comparaison compréhensible entre les différentes bases d'écorce.

6.3.2 Classification d'écorce avec ResNet34

6.3.2.1 ResNet34 entièrement entraîné

Le Tableau 6.29 présente les résultats de classification avec le modèle ResNet34 entièrement entraîné sur chacune des bases d'écorce. On peut établir les remarques suivantes à partir du Tableau 6.29 :

Les taux de classification retournés par ResNet34 entièrement entraîné à partir de poids aléatoires sont fortement corrélés avec la taille de la base d'écorce en question. En effet, on voit que le taux de classification a tendance à augmenter en fonction du nombre d'images de la base. Par exemple, moins de 50% sur la base Trunk12 avec 393 images, et 73.41% sur la base BarkNet avec plus de 23000 images. Cependant ceci n'est pas vérifié pour la base Bark101 à cause du nombre élevé de classes avec peu d'images par espèce (par fois deux images par espèce). Ceci limite l'apprentissage des caractéristiques par classe.

Par ailleurs, on trouve aussi qu'avoir une grande base avec un nombre élevé d'images nécessite plusieurs epoch pour que le modèle d'apprentissage converge, comme le cas des bases Bark101 et BarkNet. Avec les bases Trunk12 et AFF, le modèle tend vers un sur apprentissage avec un nombre élevé d'epoch.

Au final, les scores de classification avec ResNet34 restent relativement faibles dans le contexte de classification d'écorce, compte tenu de leurs performances dans d'autres domaine de classification d'images. Ceci est essentiellement dû au fait que le problème actuel est celui de la classification de différentes classes d'un même objet naturel, i.e., l'écorce, et aussi au fait qu'on ne dispose pas pour autant d'un très grand nombre d'échantillons permettant un bon apprentissage d'attributs, d'autant plus leur généralisation sur l'ensemble de test. Par conséquent, pour la classification des espèces d'écorce, l'utilisation des CNN entièrement entraînés restent limitée par la taille des bases d'écorce.

6.3.2.2 ResNet34 pré-entraîné

Le Tableau 6.30 présente les résultats de classification avec le modèle ResNet34 pré-entraîné et réadapté sur chacune des bases d'écorce. À partir du Tableau 6.30, on peut voir une amélioration importante des scores de classification avec un réseau pré-entraîné sur toutes les bases d'écorce, et ceux par rapport aux résultats affichés dans le Tableau 6.29. Par exemple, on note un écart de +41.02% pour Trunk12, +23.26% pour AFF, et +36.21% pour Bark101, et un faible écart de +06.21% sur la grande base de BarkNet. Ces augmentations des scores de classification, notamment sur les petites bases, comme Trunk12, AFF et Bark101, sont à la base dues au fait que l'entraînement du réseau ne commence pas avec des poids aléatoires, mais bénéficie de la généralisation des poids appris sur une très grande base d'objets génériques comme *ImageNet* (plus de 1.2 million d'images pour l'entraînement). Il est aussi intéressant de reporter le cas de la base Bark101 où le score est trois fois plus meilleur avec un transfert d'apprentissage.

Les scores obtenus dans le contexte de classification des espèces d'écorce montrent l'intérêt d'avoir un très grand ensemble pour l'apprentissage du réseau de neurones à convolution. Néanmoins, le transfert d'apprentissage reste une solution intéressante qui permet de bénéficier des performances des CNN, même avec des bases très différentes, comme le cas ici des images d'écorce et les images de *ImageNet*.

6.3.3 Comparaison entre SRBP, SMBP et ResNet34

Afin d'établir une comparaison significative avec nos descripteurs des macro-motifs binaires statistiques, nous procédons comme suit : pour une base d'images donnée, on passe les images dans le modèle ResNet34 et on extrait le vecteur descripteur à partir de la dernière couche (*Fully Connected*). Pour la classification, on utilise les classifieurs 1 – NN (avec L1) et SVM_{lin}. Pour les descripteurs SRBP et SMBP nous présentons les scores de classification sur des images RGB.

6.3.3.1 ResNet34 entièrement entraîné

Le Tableau 6.31 montre les résultats obtenus avec un ResNet34 entièrement entraîné sur chacune des bases d'écorce.

Il est clair qu'à partir du Tableau 6.31 nos descripteurs SRBP et SMBP retournent les scores les plus élevés sur les bases d'écorce Trunk12, AFF et Bark101 en comparaison des descripteurs d'attributs ResNet34, et ceux avec les différentes configurations de l'espace d'échelle, *circ* et *ring*, et du voisinage, *AllN* et *OneN*. Par exemple, on trouve un écart respectif de +29.23%, +20.55%, +10.58% entre le taux de classification de notre meilleur descripteur à macro-motifs statistiques et le descripteur ResNet34 pour les bases Trunk12, AFF, et Bark101 avec le classifieur SVM_{lin}. Ceci est en majeure partie influencé par la taille de l'ensemble d'apprentissage sur ces bases avec un nombre total faible d'images d'écorce.

Cependant, sur la base BarkNet, nos descripteurs retournent des scores comparables, où le descripteur d'attributs ResNet34 dépasse seulement par 2.07% le meilleur score obtenu par le descripteur *circ*-SRBP. En effet, avec un très grand nombre d'images, une approche d'extraction de caractéristiques à base d'apprentissage est plus efficace, comme le cas de la base BarkNet.

TABLEAU 6.31 – Comparaison des taux de classification de nos descripteurs SMBP et SRBP et ceux des descripteurs extraits de ResNet34 entièrement entraîné sur les bases d'écorces.

Trunk12									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	70.76	0.72	0.71	0.71	83.07	0.84	0.82	0.83
	circ-SMBP	72.82	0.73	0.72	0.72	82.56	0.82	0.82	0.82
<i>OneN</i>	circ-SRBP	70.76	0.72	0.70	0.71	78.97	0.78	0.78	0.78
	circ-SMBP	65.64	0.67	0.65	0.66	83.58	0.83	0.83	0.83
<i>AllN</i>	ring-SRBP	76.92	0.77	0.76	0.77	85.64	0.85	0.85	0.85
	ring-SMBP	77.43	0.77	0.76	0.76	84.61	0.84	0.84	0.84
<i>OneN</i>	ring-SRBP	74.87	0.76	0.74	0.75	85.12	0.86	0.84	0.85
	ring-SMBP	73.84	0.74	0.73	0.73	82.05	0.83	0.81	0.82
Attribut ResNet34		49.23	0.47	0.48	0.47	56.41	0.55	0.55	0.55
AFF									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	74.70	0.76	0.73	0.74	80.81	0.83	0.77	0.80
	circ-SMBP	78.60	0.80	0.76	0.78	81.66	0.85	0.76	0.81
<i>OneN</i>	circ-SRBP	71.81	0.72	0.70	0.71	78.60	0.74	0.72	0.73
	circ-SMBP	74.19	0.76	0.70	0.73	80.13	0.74	0.74	0.74
<i>AllN</i>	ring-SRBP	75.72	0.76	0.73	0.74	80.47	0.85	0.76	0.80
	ring-SMBP	77.41	0.80	0.74	0.77	80.81	0.76	0.75	0.76
<i>OneN</i>	ring-SRBP	75.21	0.75	0.72	0.74	82.17	0.86	0.78	0.82
	ring-SMBP	78.94	0.79	0.75	0.77	76.74	0.74	0.71	0.72
Attribut ResNet34		58.40	0.60	0.60	0.60	61.62	0.64	0.61	0.63
Bark101									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	33.82	0.24	0.23	0.23	33.74	0.18	0.16	0.17
	circ-SMBP	37.68	0.29	0.26	0.28	29.80	0.15	0.13	0.14
<i>OneN</i>	circ-SRBP	30.65	0.22	0.21	0.21	31.04	0.15	0.14	0.14
	circ-SMBP	34.28	0.25	0.24	0.25	30.73	0.16	0.14	0.15
<i>AllN</i>	ring-SRBP	34.90	0.26	0.24	0.25	32.35	0.17	0.14	0.15
	ring-SMBP	37.06	0.29	0.26	0.27	30.88	0.15	0.13	0.14
<i>OneN</i>	ring-SRBP	32.43	0.24	0.23	0.23	30.57	0.14	0.12	0.13
	ring-SMBP	36.37	0.28	0.27	0.26	30.96	0.14	0.14	0.14
Attribut ResNet34		22.39	0.17	0.16	0.17	23.16	0.17	0.13	0.15
BarkNet									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	53.46	0.52	0.50	0.51	67.50	0.69	0.67	0.68
	circ-SMBP	56.04	0.53	0.52	0.53	63.90	0.64	0.62	0.63
<i>OneN</i>	circ-SRBP	54.97	0.53	0.52	0.53	66.58	0.68	0.65	0.67
	circ-SMBP	55.30	0.53	0.51	0.52	66.60	0.67	0.66	0.66
<i>AllN</i>	ring-SRBP	52.69	0.51	0.50	0.50	67.31	0.67	0.66	0.67
	ring-SMBP	51.87	0.49	0.49	0.49	62.86	0.63	0.62	0.62
<i>OneN</i>	ring-SRBP	53.80	0.52	0.51	0.51	66.67	0.67	0.65	0.66
	ring-SMBP	53.10	0.50	0.49	0.50	64.46	0.65	0.64	0.64
Attribut ResNet34		62.72	0.62	0.62	0.62	69.57	0.69	0.68	0.69

6.3.3.2 ResNet34 pré-entraîné

Le Tableau 6.32 présente les résultats de classification obtenus avec un ResNet34 pré-entraîné et réadapté sur chacune des bases d'écorce. Les descripteurs SRBP et SMBP sont pris du Tableau 6.31 pour les variantes avec les meilleurs scores. On cherche à voir l'impact du pré-entraînement sur

TABLEAU 6.32 – Comparaison des taux de classification de nos descripteurs SMBP et SRBP et ceux des descripteurs extraits de ResNet34 pré-entraîné et réadapté sur les bases d'écorce.

Trunk12									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	ring-SRBP	76.92	0.77	0.76	0.77	85.64	0.85	0.85	0.85
	ring-SMBP	77.43	0.77	0.76	0.76	84.61	0.84	0.84	0.84
Attribut ResNet34		94.87	0.95	0.94	0.94	95.89	0.96	0.95	0.96
AFF									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>OneN</i>	ring-SRBP	75.21	0.75	0.72	0.74	82.17	0.86	0.78	0.82
	ring-SMBP	78.94	0.79	0.75	0.77	76.74	0.74	0.71	0.72
Attribut ResNet34		89.47	0.89	0.89	0.89	91.68	0.92	0.92	0.92
Bark101									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	33.82	0.24	0.23	0.23	33.74	0.18	0.16	0.17
	circ-SMBP	37.68	0.29	0.26	0.28	29.80	0.15	0.13	0.14
Attribut ResNet34		57.83	0.51	0.42	0.46	56.98	0.45	0.39	0.42
BarkNet									
Classifieurs		1 – NN				SVM _{lin}			
Descripteurs		Score (%)	Précision	Rappel	F1-score	Score (%)	Précision	Rappel	F1-score
<i>AllN</i>	circ-SRBP	53.46	0.52	0.50	0.51	67.50	0.69	0.67	0.68
	circ-SMBP	56.04	0.53	0.52	0.53	63.90	0.64	0.62	0.63
Attribut ResNet34		76.43	0.76	0.76	0.76	80.56	0.80	0.79	0.80

les attributs extraits de ResNet34.

Il est clair que les attributs extraits à partir d'un ResNet34 pré-entraîné donnent les meilleurs scores sur toutes les bases d'écorce. De plus, ces scores ont une fois de plus augmenté par rapport à ceux des attributs extraits à partir d'un ResNet34 entraîné à partir de poids aléatoires (c.f. Tableau 6.31). L'utilisation d'un modèle pré-entraîné tire avantage du transfert d'apprentissage à partir d'une grande base comme *ImageNet*, tel que les poids des couches inférieures peuvent être généralisés à d'autres applications différentes, comme la classification des espèces d'écorce. Néanmoins, malgré les résultats supérieurs des attributs ResNet34, il reste difficile d'apporter une interprétation à leurs comportements quant aux caractéristiques encodées à partir des surfaces d'écorce, contre nos descripteurs qui encodent essentiellement la primitive texture.

Toutefois, il sera intéressant d'évaluer l'application de différents modèles de réseaux de neurones à convolution pour la classification des espèces d'écorce, autres modèles que le ResNet34. Ces modèles peuvent être pré-entraînés afin de remédier au manque de grandes bases d'écorces annotées.

Au final, ces résultats justifient pour autant l'utilisation massive des réseaux de neurones à convolution pour des applications de classification des espèces végétales, notamment lors de la campagne d'évaluation PlantClef¹, à partir de différents organes végétaux.

1. <https://www.imageclef.org/>

6.4 Conclusion

Nous avons vu à travers ce chapitre que la description proposée pour les différentes surfaces de texture de *quatre* bases défiantes d'écorce, i.e., Trunk12, AFF, Bark101 et BarkNet, s'avère efficace et pertinente.

Nous avons trouvé que nos descripteurs occupent le premier rang sur les différentes bases avec des performances qui dépassent les méthodes de l'état de l'art, avec des scores d'identification de 81.17% avec ring-SMBP sur la base Trunk12, 81.38% avec circ-SMBP sur la base AFF. De plus sur les bases à grande échelle, nos descripteurs donnent des taux de classification de 37.37% en SVM_{rbf} pour ring-SRBP sur la base Bark101, avec plus de 100 espèces différentes d'écorce, et 57.27% en SVM linéaire pour circ-SRBP sur la très grande base de BarkNet. Par ailleurs, nos descripteurs distinguent mieux les espèces d'écorce avec une similarité inter-espèces, prouvant ainsi leur efficacité pour ce problème d'identification d'objets naturels.

Ce dernier est ainsi considéré comme un problème complexe, où il est difficile de caractériser la texture des écorces du fait de ses structures variables en taille et en arrangement spatial. Au final, cette étude expérimentale a permis de recommander les descripteurs à macro-motif statistique profond SMBP_m, et notamment le descripteur combiné SMBP, qui a montré une stabilité et une supériorité de performances sur les différentes bases d'écorce.

Par la suite, nous avons établi une classification d'écorce avec le modèle ResNet34 entièrement entraîné sur chacune des bases d'écorce. Les résultats obtenus montrent les faibles performances de ResNet34. De plus, une comparaison entre nos descripteurs et le vecteur d'attributs ResNet34 a montré la limite d'une approche à base d'apprentissage sur des bases de taille relativement faible, que ce soit en nombre de classes, ou en nombre d'images par classe.

Par la suite, nous avons établi une classification des espèces d'écorce à partir des deux descriptions : avec nos descripteurs, manuellement conçus, et le descripteur extrait de ResNet34, i.e., une représentation à base d'apprentissage profond.

Les vecteurs d'attributs de macro-motifs binaires statistiques proposées sont plus efficaces pour la représentation du contenu visuel des surfaces d'écorce lorsqu'il s'agit de bases d'apprentissage très petites.

Chapitre 7

Conclusions et Perspectives

Sommaire

7.1 Résumé des travaux	140
7.2 Perspectives	141

7.1 Résumé des travaux

À travers cette thèse, nous avons traité le problème d'identification et classification des espèces végétales à partir des images d'écorce. Le problème a été formalisé comme un problème de représentation et classification de texture.

Plus précisément, nous avons introduit une nouvelle représentation de la surface des écorces basée sur les motifs binaires, et inspirée de la méthode LBP (Local Binary Pattern). L'originalité de ce travail se résume en la proposition d'une représentation réduite de l'information de texture contenue dans une macrostructure avec un très grand nombre d'échelles spatiales, et cela par un seul code macro-motif binaire, à l'encontre des variantes de l'état de l'art à base de motifs binaires, qui, elles encodent la macrostructure par la concaténation de plusieurs micro-motifs par échelle.

La formalisation de la description de texture proposée dans cette thèse constitue un modèle générique, à partir duquel, plusieurs descripteurs de macro-motifs peuvent être dérivés. Ainsi, un cas spécifique à la texture d'écorce a été introduit pour encoder les macrostructures qui caractérisent la texture des écorces par des macro-motifs binaires statistiques. Ainsi, quatre familles de descripteurs statistiques sont présentées avec deux structures spatiales et deux encodages différents, à savoir, circ-SRBP_m , ring-SRBP_m pour les descripteurs radiaux, et circ-SMBP_m et ring-SMBP_m pour les descripteurs profonds, en structure circulaire *circ* et en structure à anneaux *ring*, avec différentes représentations statistiques. De plus, une complémentarité entre les différentes représentations statistiques est donnée à travers la concaténation des descripteurs statistiques pour chacun des deux encodages proposés.

Des évaluations expérimentales ont été effectuées sur quatre différentes bases d'écorce, à savoir, Trunk12, AFF, Bark101 et BarkNet, où pour chaque base, différents problèmes relatifs au nombre d'espèces, au nombre d'images, aux surfaces altérées, et au problème de similarité inter-espèces étaient présents.

En premier lieu, nous avons établi une étude comparative entre différentes variantes LBP multi-échelles afin d'examiner leurs impacts sur l'identification des espèces végétales. Nous avons montré l'intérêt d'une représentation multi-échelles pour l'identification des différentes espèces d'écorce, ainsi que leurs limites avec un large nombre d'échelles, i.e., la dimensionnalité du descripteur final, et sa performance à distinguer les différentes structures des écorces. D'où la proposition de notre solution de la représentation réduite de la macrostructure sur de larges supports spatiaux.

En second lieu, nous avons détaillé l'évaluation des descripteurs proposés. Pour faire, nous avons procédé à une évaluation paramétrique suivie d'une évaluation comparative avec l'état de l'art. Nous avons montré la supériorité de la représentation proposée en terme de score d'identification et de MAP sur les bases Trunk12 et AFF avec les valeurs respectives de 81.17%, 43.54 avec ring-SMBP , et 81.38%, 31.50 avec circ-SMBP . De plus, nos descripteurs peuvent distinguer entre des espèces d'écorce très similaires dont les images ont été prises sous différentes conditions non contrôlées. Sur les bases à grande échelle, Bark101 et BarkNet, les descripteurs des macro-motifs binaires statistiques ont une fois de plus montré leurs efficacité pour la classification d'un très grand nombre d'espèces végétales, et un large nombre d'images d'écorce, donnant ainsi les meilleurs scores sur ces bases avec 37.37% en SVM_{rbf} pour ring-SRBP sur la base Bark101, avec

plus de 100 espèces différentes d'écorce, et 57.27% en SVM linéaire pour circ-SRBP sur la très grande base de BarkNet.

Au final, en contraste avec notre conception classique pour la représentation et identification des espèces d'écorce, nous avons effectué une classification basée sur l'apprentissage profond, avec un modèle de réseaux de neurones à convolution, i.e., ResNet34. Le modèle ResNet34 était entièrement entraîné sur les bases d'écorce. Les taux de classification enregistrés ont tendance à augmenter en fonction du nombre d'images par espèce utilisées pour l'entraînement.

Ensuite, nous avons présenté un comparatif entre nos descripteurs avec ceux extraits à partir de ResNet34, afin d'établir un ordre de grandeur entre deux approches différentes de la représentation du contenu visuel, à savoir, l'approche classique représentée par nos descripteurs à macro-motifs binaires statistiques, et l'approche basée sur l'apprentissage profond avec ResNet34. Les taux de classification de nos descripteurs demeurent significativement supérieures à ceux du descripteur d'attributs ResNet34 sur les bases à faible nombre d'images d'écorce. Plus précisément, nos descripteurs à macro-motifs binaires statistiques enregistrent des écarts de +29.23%, +20.55%, +10.58% en taux de classification pour les bases Trunk12, AFE, et Bark101 avec SVM linéaire. Tandis que sur la très grande base de BarkNet, nous avons obtenu des résultats comparatifs où le ResNet34 dépasse seulement de +2.07% avec circ-SRBP. Ces résultats expérimentaux montrent l'avantage de notre approche pour la classification des différentes espèces d'écorce, en particulier, sur les bases d'écorce avec peu d'images. Cela sans l'introduction d'une phase d'apprentissage automatique des attributs comme pour le ResNet34, dont l'efficacité est moindre lorsque la base contient très peu d'images et/ou un très grand nombre d'espèces.

7.2 Perspectives

L'étude présentée dans cette thèse sur l'identification et la classification des espèces d'écorce peut être étendue selon de multiples perspectives comme suit :

- ▶ Étendre l'encodage par des descripteurs statistiques de macro-motifs binaires à un contexte adaptatif à l'échelle de la structure d'écorce. En effet, la représentation optimale de la macrostructure proposée dans cette étude est essentiellement basée sur une recherche empirique des paramètres qui régularisent sa configuration. Nous pensons qu'une représentation adaptative de l'échelle, i.e., échelle globale de l'image, ou bien échelle locale de la région de la macrostructure, permet, en plus de l'invariance au changement de la rotation locale, d'apporter une invariance à l'échelle, et améliorer les performances de l'identification de la texture d'écorce.
- ▶ Introduire un nouveau modèle CNN spécifique à la classification des espèces d'écorce en se basant sur les modèles CNN introduits sur des images de textures génériques.
- ▶ Étendre l'application de nos descripteurs de macro-motifs binaires statistiques à la description du contenu visuel des images de différents domaines spécifiques, comme des bases médicales, à savoir, des bases des protéines, des lésions dermatologiques, des cellules et des radiographies de cancer, etc. En effet, pour ce genre d'applications, nous retrouvons des particularités similaires à notre application sur les écorces, notamment, le faible nombre d'images et/ou de classes, une seule catégorie d'objet naturel à identifier, et la primitive

'texture' qui constitue le principal attribut visuel de ces images. De ce fait, il sera judicieux d'examiner l'apport de notre description dans différents domaines.

- Généraliser l'application de nos descripteurs statistiques des macro-motifs binaires statistiques sur des bases de textures génériques plus ou moins défiants. En effet, nous avons réalisé une expérimentation préliminaire sur la base de texture générique Outex_TC_00010 avec nos descripteurs, et nous avons trouvé des résultats intéressants dont le score avoisine l'état de l'art sur cette base. De ce fait, une perspective relative à l'applicabilité de notre approche de représentation de la texture reste à étudier sur les différentes bases de textures génériques, comme CURET, KTHTips, KTHTips2a/2b, UIUC, etc.

Références

- AHONEN, T., A. HADID et M. PIETIKAINEN. 2006, «Face description with local binary patterns : Application to face recognition», *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, n° 12, p. 2037–2041. [30](#)
- AHONEN, T., J. MATAS, C. HE et M. PIETIKÄINEN. 2009, «Rotation invariant image description with local binary pattern histogram fourier features», dans *Scandinavian Conference on Image Analysis*, Springer, p. 61–70. [35](#), [36](#)
- AIBINU, A. M., M.-J. E. SALAMI, A. A. SHAFIE, N. HAZALI et N. TERMIDZI. 2011, «Automatic fruits identification system using hybrid technique», dans *Electronic Design, Test and Application (DELTA), 2011 Sixth IEEE International Symposium on*, IEEE, p. 217–221. [14](#)
- BAKIĆ, V., S. MOUINE, S. OUERTANI-LITAYEM, A. VERROUST-BLONDET, I. YAHIAOUI, H. GOËAU et A. JOLY. 2013, «Inria's participation at imageclef 2013 plant identification task», dans *Inria's participation at imageclef 2013 plant identification task*, p. –. [15](#), [19](#), [24](#), [37](#), [38](#)
- BAY, H., T. TUYTELAARS et L. VAN GOOL. 2006, «Surf : Speeded up robust features», dans *European conference on computer vision*, Springer, p. 404–417. [23](#), [24](#), [61](#)
- BELHUMEUR, P. N., D. CHEN, S. FEINER, D. W. JACOBS, W. J. KRESS, H. LING, I. LOPEZ, R. RAMAMOORTHI, S. SHEOREY, S. WHITE et collab.. 2008, «Searching the world's herbaria : A system for visual identification of plant species», dans *European Conference on Computer Vision*, Springer, p. 116–129. [2](#), [3](#), [14](#)
- BENGIO, Y., A. COURVILLE et P. VINCENT. 2013, «Representation learning : A review and new perspectives», *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, n° 8, p. 1798–1828. [39](#)
- BERTRAND, S., G. CERUTTI et L. TOUGNE. 2017, «Bark recognition to improve leaf-based classification in didactic tree species identification.», dans *Computer Vision Theory and Applications, International Conference on*, p. 435–442. [ix](#), [4](#), [14](#), [15](#), [21](#), [22](#), [37](#), [38](#), [111](#), [118](#)
- BLAANCO, L. J., C. M. TRAVIESO, J. M. QUINTEIRO, P. V. HERNANDEZ, M. K. DUTTA et A. SINGH. 2016, «A bark recognition algorithm for plant classification using a least square support vector machine», dans *Contemporary Computing (IC3), 2016 Ninth International Conference on*, IEEE, p. 1–5. [15](#), [19](#), [38](#)
- BO, L., X. REN et D. FOX. 2010, «Kernel descriptors for visual recognition», dans *Advances in neural information processing systems*, p. 244–252. [19](#)
- BRESSANE, A., J. A. ROVEDA et A. C. MARTINS. 2015a, «Statistical analysis of texture in trunk images for biometric identification of tree species», *Environmental monitoring and assessment*, vol. 187, n° 4, p. 212. [17](#), [38](#)

- BRESSANE, A., J. A. F. ROVEDA, A. C. GERMANO et S. R. M. M. R. MARTINS. 2016, «Pattern recognition based on soft boundaries : a proposal applied to tree species identification from texture in trunk images», *Congresso Brasileiro de Sistemas*, p. –. [15](#), [17](#), [36](#), [38](#), [89](#)
- BRESSANE, A., J. A. F. ROVEDA et A. C. G. MARTINS. 2015b, «Pattern recognition in trunk images based on co-occurrence descriptors : A proposal applied to tree species identification», dans *Computational Intelligence (LA-CCI), 2015 Latin America Congress on*, IEEE, p. 1–6. [15](#), [17](#), [36](#), [38](#), [89](#)
- CANNY, J. 1987, «A computational approach to edge detection», dans *Readings in Computer Vision*, Elsevier, p. 184–203. [21](#)
- CARPENTIER, M., P. GIGUERE et J. GAUDREULT. 2018, «Tree species identification from bark images using convolutional neural networks», dans *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, p. 1075–1081. [4](#), [41](#), [86](#), [87](#), [88](#), [94](#), [129](#), [133](#)
- CERUTTI, G., L. TOUGNE, J. MILLE, A. VACAVANT et D. COQUIN. 2013, «Understanding leaves in natural images—a model-based approach for tree species identification», *Computer Vision and Image Understanding*, vol. 117, n° 10, p. 1482–1501. [3](#), [44](#)
- CHANG, C.-C. et C.-J. LIN. 2011, «Libsvm : A library for support vector machines», *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, n° 3, p. 27. [89](#)
- CHEN, Q., M. ABEDINI, R. GARNAVI et X. LIANG. 2014, «Ibm research australia at lifeclef2014 : Plant identification task.», dans *CLEF (Working Notes)*, p. 693–704. [15](#), [24](#), [37](#), [38](#), [44](#)
- CHI, Z., L. HOUQIANG et W. CHAO. 2003, «Plant species recognition based on bark patterns using novel gabor filter banks», dans *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*, vol. 2, IEEE, p. 1035–1038. [4](#), [9](#), [15](#), [21](#), [37](#), [89](#)
- CIMPOI, M., S. MAJI, I. KOKKINOS et A. VEDALDI. 2016, «Deep filter banks for texture recognition, description, and segmentation», *International Journal of Computer Vision*, vol. 118, n° 1, p. 65–94. [40](#)
- COPE, J. S., D. CORNEY, J. Y. CLARK, P. REMAGNINO et P. WILKIN. 2012, «Plant species identification using digital morphometrics : A review», *Expert Systems with Applications*, vol. 39, n° 8, p. 7562–7573. [3](#), [14](#)
- CSURKA, G., C. DANCE, L. FAN, J. WILLAMOWSKI et C. BRAY. 2004, «Visual categorization with bags of keypoints», dans *Workshop on statistical learning in computer vision, ECCV*, vol. 1, Prague, p. 1–22. [23](#), [62](#)
- DO, M. N. et M. VETTERLI. 2005, «The contourlet transform : an efficient directional multiresolution image representation», *IEEE Transactions on image processing*, vol. 14, n° 12, p. 2091–2106. [22](#)
- FEKRI-ERSHAD, S. 2020, «Bark texture classification using improved local ternary patterns and multilayer neural network», *Expert Systems with Applications*, p. 113 509. [14](#), [38](#)

- FERNÁNDEZ, A., M. X. ÁLVAREZ et F. BIANCONI. 2011, «Image classification with binary gradient contours», *Optics and Lasers in Engineering*, vol. 49, n° 9-10, p. 1177–1184. [33](#), [35](#)
- FU, X. et W. WEI. 2008, «Centralized binary patterns embedded with image euclidean distance for facial expression recognition», dans *Natural Computation, 2008. ICNC'08. Fourth International Conference on*, vol. 4, IEEE, p. 115–119. [33](#), [35](#)
- GODET, J. 2012, *Guide des écorces des arbres d'Europe : reconnaître et comparer les espèces*, Guides du naturaliste, Delachaux et Niestlé, ISBN 9782603018491. URL <https://books.google.fr/books?id=H-wUkgEACAAJ>. [6](#), [10](#), [45](#)
- GOËAU, H., P. BONNET, J. BARBE, V. BAKIC, A. JOLY, J.-F. MOLINO, D. BARTHELEMY et N. BOUJEMAA. 2012, «Multi-organ plant identification», dans *Proceedings of the 1st ACM international workshop on Multimedia analysis for ecological data*, ACM, p. 41–44. [14](#), [15](#), [37](#)
- GOËAU, H., P. BONNET et A. JOLY. 2015, «Lifeclef plant identification task 2015», dans *Working Notes of CLEF 2015 - Conference and Labs of the Evaluation forum Toulouse, France, September 8-11, 2015.*, CEUR-WS, p. 0–0. [14](#), [44](#)
- GOËAU, H., P. BONNET, A. JOLY, N. BOUJEMAA, D. BARTHELEMY, J.-F. MOLINO, P. BIRNBAUM, E. MOUYSSET et M. PICARD. 2011, «The imageclef 2011 plant images classification task», dans *ImageCLEF 2011*, p. 0–0. [3](#)
- GOËAU, H., A. JOLY, P. BONNET, V. BAKIC, D. BARTHÉLÉMY, N. BOUJEMAA et J.-F. MOLINO. 2013, «The imageclef plant identification task 2013», dans *Proceedings of the 2nd ACM international workshop on Multimedia analysis for ecological data*, ACM, p. 23–28. [14](#)
- GOËAU, H., A. JOLY, P. BONNET, S. SELMI, J.-F. MOLINO, D. BARTHÉLÉMY et N. BOUJEMAA. 2014a, «Lifeclef plant identification task 2014», dans *CLEF2014 Working Notes. Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014*, CEUR-WS, p. 598–615. [14](#)
- GOËAU, H., A. JOLY, I. YAHIAOUI, V. BAKIĆ, A. VERROUST-BLONDET, P. BONNET, D. BARTHÉLÉMY, N. BOUJEMAA et J.-F. MOLINO. 2014b, «Plantnet participation at lifeclef2014 plant identification task», dans *CLEF2014 Working Notes. Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014*, CEUR-WS, p. 724–737. [15](#), [19](#), [24](#), [37](#), [38](#)
- GOODFELLOW, I., Y. BENGIO et A. COURVILLE. 2016, *Deep learning*, MIT press. [39](#)
- GUO, Z., L. ZHANG et D. ZHANG. 2010a, «A completed modeling of local binary pattern operator for texture classification», *IEEE Transactions on Image Processing*, vol. 19, n° 6, p. 1657–1663. [32](#), [34](#), [35](#), [36](#), [58](#), [83](#), [111](#), [118](#)
- GUO, Z., L. ZHANG et D. ZHANG. 2010b, «Rotation invariant texture classification using lbp variance (LBPV) with global matching», *Pattern recognition*, vol. 43, n° 3, p. 706–719. [35](#), [36](#), [111](#), [117](#), [118](#)
- GUO, Z., L. ZHANG, D. ZHANG et S. ZHANG. 2010c, «Rotation invariant texture classification using adaptive lbp with directional statistical features», dans *Image Processing (ICIP), 2010 17th IEEE International Conference on*, IEEE, p. 285–288. [35](#)

- HAFIANE, A., G. SEETHARAMAN et B. ZAVIDOVIQUE. 2007, «Median binary pattern for textures classification», dans *International Conference Image Analysis and Recognition*, Springer, p. 387–398. [33](#), [35](#)
- HARALICK, R. M., K. SHANMUGAM et collab.. 1973, «Textural features for image classification», *IEEE Transactions on systems, man, and cybernetics*, p. 610–621. [16](#), [17](#)
- HARRIS, C. et M. STEPHENS. 1988, «A combined corner and edge detector.», dans *Alvey vision conference*, vol. 15, Citeseer, p. 10–5244. [24](#), [61](#)
- HE, K., X. ZHANG, S. REN et J. SUN. 2016, «Deep residual learning for image recognition», dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 770–778. [39](#), [94](#)
- HEIKKILÄ, M., M. PIETIKÄINEN et C. SCHMID. 2006, «Description of interest regions with center-symmetric local binary patterns», dans *Computer vision, graphics and image processing*, Springer, p. 58–69. [33](#), [35](#)
- HUANG, D., C. SHAN, M. ARDABILIAN, Y. WANG et L. CHEN. 2011, «Local binary patterns and its application to facial image analysis : a survey», *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, n° 6, p. 765–781. [25](#)
- HUANG, G., Z. LIU, L. VAN DER MAATEN et K. Q. WEINBERGER. 2017, «Densely connected convolutional networks», dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 4700–4708. [39](#)
- HUANG, Z.-K. 2006, «Bark classification using rbpnn based on both color and texture feature», *International Journal of Computer Science and Network Security*, vol. 6, n° 11, p. 100–103. [15](#), [20](#), [36](#), [38](#), [89](#)
- HUANG, Z.-K., D.-S. HUANG, M. R. LYU et T.-M. LOK. 2006a, «Classification based on gabor filter using rbpnn classification», dans *Computational Intelligence and Security, 2006 International Conference on*, vol. 1, IEEE, p. 759–762. [21](#)
- HUANG, Z.-K., D.-S. HUANG et Z.-H. QUAN. 2006b, «Bark classification using RBPNN based on gabor filter in different color space», dans *Information Acquisition, 2006 IEEE International Conference on*, IEEE, p. 946–950. [15](#), [21](#), [37](#), [38](#), [89](#), [111](#), [118](#), [125](#), [131](#)
- HUANG, Z.-K., Z.-H. QUAN et J.-X. DU. 2006c, «Bark classification based on contourlet filter features using rbpnn», dans *International Conference on Intelligent Computing*, Springer, p. 1121–1126. [9](#), [14](#), [15](#), [22](#), [36](#), [38](#), [44](#), [89](#)
- HUANG, Z.-K., C.-H. ZHENG, J.-X. DU et Y.-Y. WAN. 2006d, «Bark classification based on textural features using artificial neural networks», dans *International Symposium on Neural Networks*, Springer, p. 355–360. [9](#), [15](#), [17](#), [36](#), [38](#), [89](#), [111](#), [118](#), [125](#), [131](#)
- JIN, H., Q. LIU, H. LU et X. TONG. 2004, «Face detection using improved lbp under bayesian framework», dans *Image and Graphics (ICIG'04), Third International Conference on*, IEEE, p. 306–309. [33](#), [35](#)

- KE-CHEN, S., Y. YUN-HUI, C. WEN-HUI et X. ZHANG. 2013, «Research and perspective on local binary pattern», *Acta Automatica Sinica*, vol. 39, n° 6, p. 730–744. [25](#)
- KHELLAH, F. M. 2011, «Texture classification using dominant neighborhood structure», *IEEE Transactions on image processing*, vol. 20, n° 11, p. 3270–3279. [35](#)
- KIM, S.-J., B.-W. KIM et D.-P. KIM. 2011, «Tree recognition for landscape using by combination of features of its leaf, flower and bark», dans *SICE Annual Conference (SICE), 2011 Proceedings of*, IEEE, p. 1147–1151. [14](#), [15](#), [20](#), [37](#), [89](#)
- KINGSBURY, N. 2001, «Complex wavelets for shift invariant analysis and filtering of signals», *Applied and computational harmonic analysis*, vol. 10, n° 3, p. 234–253. [22](#)
- KRIZHEVSKY, A., I. SUTSKEVER et G. E. HINTON. 2012, «Imagenet classification with deep convolutional neural networks», dans *Advances in neural information processing systems*, p. 1097–1105. [39](#)
- KUMAR, N., P. N. BELHUMEUR, A. BISWAS, D. W. JACOBS, W. J. KRESS, I. C. LOPEZ et J. V. SOARES. 2012, «Leafsnap : A computer vision system for automatic plant species identification», dans *European conference on computer vision*, Springer, p. 502–516. [3](#), [44](#)
- LAMEI, M., L. YAMIN et Z. HUANRONG. 2010, «Method of plant texture image recognition based on genetic programming», dans *Computer Design and Applications (ICDDA), 2010 International Conference on*, vol. 1, IEEE, p. VI–370. [15](#), [17](#), [89](#)
- LANGFORD, E. 2006, «Quartiles in elementary statistics», *Journal of Statistics Education*, vol. 14, n° 3. [72](#)
- LE, T.-L., N.-D. DUONG, H. VU et T. T.-N. NGUYEN. 2015, «Mica at lifeclef 2015 : Multi-organ plant identification.», dans *CLEF (Working Notes)*, p. -. [15](#), [19](#), [38](#), [44](#)
- LE BRAS, G., M. PIGNAL, M. L. JEANSON, S. MULLER, C. AUPIC, B. CARRÉ, G. FLAMENT, M. GAUDEUL, C. GONÇALVES, V. R. INVERNÓN et collab.. 2017, «The french muséum national d’histoire naturelle vascular plant herbarium collection dataset», *Scientific data*, vol. 4, n° 1, p. 1–16. [2](#)
- LECUN, Y., Y. BENGIO et G. HINTON. 2015, «Deep learning», *nature*, vol. 521, n° 7553, p. 436–444. [39](#)
- LEUNG, T. et J. MALIK. 2001, «Representing and recognizing the visual appearance of materials using three-dimensional textons», *International journal of computer vision*, vol. 43, n° 1, p. 29–44. [29](#)
- LIAO, S., X. ZHU, Z. LEI, L. ZHANG et S. Z. LI. 2007, «Learning multi-scale block local binary patterns for face recognition», dans *International Conference on Biometrics*, Springer, p. 828–837. [ix](#), [25](#), [26](#), [27](#), [35](#), [56](#), [68](#), [69](#), [95](#), [96](#), [97](#), [98](#)
- LIU, L., J. CHEN, P. FIEGUTH, G. ZHAO, R. CHELLAPPA et M. PIETIKÄINEN. 2019, «From bow to cnn : Two decades of texture representation for texture classification», *International Journal of Computer Vision*, vol. 127, n° 1, p. 74–109. [15](#)

- LIU, L., P. FIEGUTH, Y. GUO, X. WANG et M. PIETIKÄINEN. 2017, «Local binary features for texture classification : taxonomy and experimental study», *Pattern Recognition*, vol. 62, p. 135–160. [25](#), [30](#), [45](#), [53](#), [68](#)
- LIU, L., P. FIEGUTH, G. ZHAO, M. PIETIKÄINEN et D. HU. 2016a, «Extended local binary patterns for face recognition», *Information Sciences*, vol. 358, p. 56–72. [33](#), [34](#), [35](#), [58](#), [111](#), [118](#)
- LIU, L., S. LAO, P. W. FIEGUTH, Y. GUO, X. WANG et M. PIETIKÄINEN. 2016b, «Median robust extended local binary pattern for texture classification», *IEEE Transactions on Image Processing*, vol. 25, n° 3, p. 1368–1381. [xi](#), [31](#), [33](#), [34](#), [35](#), [56](#), [58](#), [69](#), [111](#), [116](#), [117](#), [118](#), [119](#), [132](#)
- LIU, L., Y. LONG, P. W. FIEGUTH, S. LAO et G. ZHAO. 2014, «Brint : binary rotation invariant and noise tolerant texture classification», *IEEE transactions on Image Processing*, vol. 23, n° 7, p. 3071–3084. [31](#), [32](#), [34](#), [35](#), [58](#), [68](#), [83](#), [111](#), [118](#)
- LIU, L., L. ZHAO, Y. LONG, G. KUANG et P. FIEGUTH. 2012a, «Extended local binary patterns for texture classification», *Image and Vision Computing*, vol. 30, n° 2, p. 86–99. [33](#), [34](#), [35](#), [57](#)
- LIU, L., L. ZHAO, Y. LONG, G. KUANG et P. FIEGUTH. 2012b, «Extended local binary patterns for texture classification», *Image and Vision Computing*, vol. 30, n° 2, p. 86–99. [111](#), [117](#), [118](#)
- LOWE, D. G. 2004, «Distinctive image features from scale-invariant keypoints», *International journal of computer vision*, vol. 60, n° 2, p. 91–110. [ix](#), [23](#), [61](#)
- MÄENPÄÄ, T. et M. PIETIKÄINEN. 2003, «Multi-scale binary patterns for texture analysis», dans *Scandinavian Conference on Image Analysis*, Springer, p. 885–892. [ix](#), [25](#), [26](#), [27](#), [31](#), [35](#), [55](#), [56](#), [68](#), [69](#), [95](#), [96](#), [97](#), [98](#), [111](#), [118](#)
- MALLAT, S. G. 1989, «A theory for multiresolution signal decomposition : the wavelet representation», *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, n° 7, p. 674–693. [20](#)
- MANJUNATH, B. S. et W.-Y. MA. 1996, «Texture features for browsing and retrieval of image data», *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, n° 8, p. 837–842. [21](#)
- MIKOLAJCZYK, K. et C. SCHMID. 2002, «An affine invariant interest point detector», dans *European conference on computer vision*, Springer, p. 128–142. [23](#)
- MIKOLAJCZYK, K. et C. SCHMID. 2005, «A performance evaluation of local descriptors», *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, n° 10, p. 1615–1630. [23](#)
- MOUINE, S., I. YAHIAOUI et A. VERROUST-BLONDET. 2013a, «A shape-based approach for leaf classification using multiscaletriangular representation», dans *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, ACM, p. 127–134. [14](#)
- MOUINE, S., I. YAHIAOUI, A. VERROUST-BLONDET, L. JOYEUX, S. SELMI et H. GOËAU. 2013b, «An android application for leaf-based plant identification», dans *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, ACM, p. 309–310. [3](#)

- MURALA, S., R. MAHESHWARI et R. BALASUBRAMANIAN. 2012, «Local tetra patterns : a new feature descriptor for content-based image retrieval», *IEEE Transactions on Image Processing*, vol. 21, n° 5, p. 2874–2886. [30](#)
- MZOUGH, O., I. YAHIAOUI, N. BOUJEMAA et E. ZAGROUBA. 2016, «Semantic-based automatic structuring of leaf images for advanced plant species identification», *Multimedia Tools and Applications*, vol. 75, n° 3, p. 1615–1646. [14](#), [44](#)
- NAKAYAMA, H. 2013, «Nlab-utokyo at imageclef 2013 plant identification task.», dans *CLEF (Working Notes)*, Citeseer, p. –. [15](#), [24](#), [37](#), [38](#)
- NANNI, L., A. LUMINI et S. BRAHNAM. 2010, «Local binary patterns variants as texture descriptors for medical image analysis», *Artificial intelligence in medicine*, vol. 49, n° 2, p. 117–125. [30](#), [33](#), [35](#), [59](#)
- NANNI, L., A. LUMINI et S. BRAHNAM. 2012, «Survey on lbp based texture descriptors for image classification», *Expert Systems with Applications*, vol. 39, n° 3, p. 3634–3641. [25](#), [30](#)
- NGUYEN, T. P., N.-S. VU et A. MANZANERA. 2016, «Statistical binary patterns for rotational invariant texture classification», *Neurocomputing*, vol. 173, p. 1565–1577. [32](#), [35](#), [75](#), [83](#), [111](#), [117](#), [118](#)
- NILSBACK, M.-E. et A. ZISSERMAN. 2010, «Delving deeper into the whorl of flower segmentation», *Image and Vision Computing*, vol. 28, n° 6, p. 1049–1062. [14](#)
- OJALA, T., M. PIETIKÄINEN et D. HARWOOD. 1996, «A comparative study of texture measures with classification based on featured distributions», *Pattern recognition*, vol. 29, n° 1, p. 51–59. [18](#), [35](#)
- OJALA, T., M. PIETIKÄINEN et T. MAENPÄÄ. 2002, «Multiresolution gray-scale and rotation invariant texture classification with local binary patterns», *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, n° 7, p. 971–987. [18](#), [19](#), [25](#), [35](#), [36](#), [42](#), [52](#), [53](#), [54](#), [57](#), [58](#), [64](#), [65](#), [68](#), [91](#), [95](#), [97](#), [111](#), [118](#), [125](#), [131](#)
- OTHMANI, A., L. F. L. Y. VOON, C. STOLZ et A. PIBOULE. 2013, «Single tree species classification from terrestrial laser scanning data for forest inventory», *Pattern Recognition Letters*, vol. 34, n° 16, p. 2144–2150. [15](#), [22](#), [38](#)
- PAN, S. J. et Q. YANG. 2009, «A survey on transfer learning», *IEEE Transactions on knowledge and data engineering*, vol. 22, n° 10, p. 1345–1359. [40](#)
- QI, X., Y. QIAO, C.-G. LI et J. GUO. 2013, «Multi-scale joint encoding of local binary patterns for texture and material classification.», dans *BMVC*. [57](#), [111](#), [117](#), [118](#)
- QI, X., R. XIAO, C.-G. LI, Y. QIAO, J. GUO et X. TANG. 2014, «Pairwise rotation invariant co-occurrence local binary pattern», *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, n° 11, p. 2199–2213. [57](#), [111](#), [117](#), [118](#)

- QIAN, X., X.-S. HUA, P. CHEN et L. KE. 2011, «Plbp : An effective local binary patterns texture descriptor with pyramid representation», *Pattern Recognition*, vol. 44, n° 10-11, p. 2502–2515. [ix](#), [25](#), [26](#), [27](#), [35](#), [95](#), [96](#), [98](#), [99](#), [111](#), [118](#)
- RATAJCZAK, R., S. BERTRAND, C. CRISPIM-JUNIOR et L. TOUGNE. 2019, «Efficient bark recognition in the wild», dans *International Conference on Computer Vision Theory and Applications (VISAPP 2019)*. [xiii](#), [xiv](#), [86](#), [87](#), [89](#), [111](#), [118](#), [125](#)
- REMEŠ, V. et M. HAINDL. 2019, «Bark recognition using novel rotationally invariant multispectral textural features», *Pattern Recognition Letters*, vol. 125, p. 612–617. [14](#)
- RUSSAKOVSKY, O., J. DENG, H. SU, J. KRAUSE, S. SATHEESH, S. MA, Z. HUANG, A. KARPATHY, A. KHOSLA, M. BERNSTEIN et collab.. 2015, «Imagenet large scale visual recognition challenge», *International journal of computer vision*, vol. 115, n° 3, p. 211–252. [39](#)
- SCHMID, C. 2001, «Constructing models for content-based image retrieval», dans *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 2, IEEE, p. II–II. [29](#)
- SEELAND, M., M. RZANNY, N. ALAQRAA, J. WÄLDCHEN et P. MÄDER. 2017, «Plant species classification using flower images—a comparative study of local feature representations», *PloS one*, vol. 12, n° 2, p. e0170629. [14](#)
- SIMONYAN, K. et A. ZISSERMAN. 2014, «Very deep convolutional networks for large-scale image recognition», *arXiv preprint arXiv :1409.1556*. [39](#)
- SIXTA, T. 2011, *Image and Video-based Recognition of Natural Objects*, Czech Technical University in Prague. [15](#), [25](#), [26](#), [38](#), [69](#), [86](#), [111](#), [118](#)
- SONG, J., Z. CHI, J. LIU et H. FU. 2004, «Bark classification by combining grayscale and binary texture features», dans *Intelligent Multimedia, Video and Speech Processing, 2004. Proceedings of 2004 International Symposium on*, IEEE, p. 450–453. [15](#), [21](#), [38](#), [89](#)
- ŠULC, M. et J. MATAS. 2013, «Kernel-mapped histograms of multi-scale lbps for tree bark recognition», dans *Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference of*, IEEE, p. 82–87. [4](#), [14](#), [15](#), [25](#), [26](#), [38](#), [45](#), [56](#), [69](#), [86](#), [87](#), [95](#), [96](#), [98](#), [111](#), [118](#), [125](#), [129](#), [131](#)
- SULC, M. et J. MATAS. 2014, «Fast features invariant to rotation and scale of texture», dans *European Conference on Computer Vision*, Springer, p. 47–62. [15](#), [25](#), [38](#), [45](#), [86](#)
- ŠULC, M. et J. MATAS. 2017, «Fine-grained recognition of plants from images», *Plant methods*, vol. 13, n° 1, p. 115. [4](#), [9](#), [14](#), [15](#), [25](#), [26](#), [38](#), [45](#), [86](#)
- SVAB, M. 2014, *Computer-vision-based tree trunk recognition*, Bsc thesis, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Slovenia. [86](#)

- SZEGEDY, C., W. LIU, Y. JIA, P. SERMANET, S. REED, D. ANGUELOV, D. ERHAN, V. VANHOUCKE et A. RABINOVICH. 2015, «Going deeper with convolutions», dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 1–9. [39](#)
- TAN, X. et B. TRIGGS. 2010, «Enhanced local texture feature sets for face recognition under difficult lighting conditions», *IEEE transactions on image processing*, vol. 19, n° 6, p. 1635–1650. [30](#), [33](#), [35](#), [58](#), [59](#)
- TREFNÝ, J. et J. MATAS. 2010, «Extended set of local binary patterns for rapid object detection», dans *Computer Vision Winter Workshop*, p. 1–7. [33](#), [35](#)
- TSCHIEDEL GUEDES, S. et D. BORGES. 2018, «Complex wavelet features for bark texture classification», –, p. –. [15](#), [22](#), [37](#), [89](#)
- TUCERYAN, M. et A. K. JAIN. 1993, «Texture analysis», dans *Handbook of pattern recognition and computer vision*, World Scientific, p. 235–276. [ix](#), [15](#), [16](#)
- VARMA, M. et A. ZISSERMAN. 2005, «A statistical approach to texture classification from single images», *International journal of computer vision*, vol. 62, n° 1-2, p. 61–81. [ix](#), [29](#), [30](#), [45](#)
- VARMA, M. et A. ZISSERMAN. 2009, «A statistical approach to material classification using image patch exemplars», *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, n° 11, p. 2032–2047. [29](#), [45](#)
- VAUCHER, H. 1993, *Guide des écorces*, Delachaux et Niestlé, ISBN 9782603008508. URL <https://books.google.fr/books?id=o3vuHAAACAAJ>. [6](#), [8](#), [10](#), [14](#), [45](#)
- VIOLA, P. et M. JONES. 2001, «Robust real-time object detection», dans *International Journal of Computer Vision*, p. 34–47. [25](#)
- VU, N.-S. et A. CAPLIER. 2012, «Enhanced patterns of oriented edge magnitudes for face recognition and image matching», *IEEE Transactions on Image Processing*, vol. 21, n° 3, p. 1352–1365. [32](#)
- WÄLDCHEN, J. et P. MÄDER. 2018, «Plant species identification using computer vision techniques : A systematic literature review», *Archives of Computational Methods in Engineering*, vol. 25, n° 2, p. 507–543. [3](#)
- WAN, Y.-Y., J.-X. DU, D.-S. HUANG, Z. CHI, Y.-M. CHEUNG, X.-F. WANG et G.-J. ZHANG. 2004, «Bark texture feature extraction based on statistical texture analysis», dans *Intelligent Multimedia, Video and Speech Processing, 2004. Proceedings of 2004 International Symposium on*, IEEE, p. 482–485. [9](#), [14](#), [15](#), [17](#), [36](#), [38](#), [89](#), [111](#), [118](#), [125](#), [131](#)
- WANG, K., C.-E. BICHOT, Y. LI et B. LI. 2017, «Local binary circumferential and radial derivative pattern for texture classification», *Pattern Recognition*, vol. 67, p. 213–229. [33](#), [35](#)
- WANG, K., C.-E. BICHOT, C. ZHU et B. LI. 2013, «Pixel to patch sampling structure and local neighboring intensity relationship patterns for texture classification», *IEEE Signal Processing Letters*, vol. 20, n° 9, p. 853–856. [31](#), [33](#), [35](#), [56](#), [68](#), [111](#), [117](#), [118](#)

- WANG, X. et Z. WANG. 2013, «A novel method for image retrieval based on structure elements' descriptor», *Journal of Visual Communication and Image Representation*, vol. 24, n° 1, p. 63–74. [30](#)
- WENDEL, A., S. STERNIG et M. GODEC. 2011, «Automated identification of tree species from images of the bark, leaves and needles», dans *16th Computer Vision Winter Workshop*, Citeseer, p. 67. [4](#), [9](#), [14](#), [15](#), [24](#), [38](#), [86](#), [118](#)
- WILLIS, K. J. 2017, *State of the World's Plants 2017*, London (UK) : Royal Botanic Gardens, Kew. URL <https://stateoftheworldsplants.org/>. [2](#)
- WOJTECH, M. 2011, *Bark : A Field Guide to Trees of the Northeast*, University Press of New England, ISBN 9781584658528. URL <https://books.google.fr/books?id=OSWqcQAACAAJ>. [ix](#), [6](#), [7](#), [8](#), [10](#), [14](#), [16](#), [45](#)
- WOLF, L., T. HASSNER et Y. TAIGMAN. 2008, «Descriptor based methods in the wild», dans *Workshop on faces in real-life images : Detection, alignment, and recognition*, p. –. [31](#), [35](#), [56](#), [69](#), [111](#), [118](#)
- ZAWBAA, H. M., M. ABBASS, M. HAZMAN et A. E. HASSENIAN. 2014, «Automatic fruit image recognition system based on shape and color features», dans *International Conference on Advanced Machine Learning Technologies and Applications*, Springer, p. 278–290. [14](#)
- ZHANG, B., Y. GAO, S. ZHAO et J. LIU. 2010, «Local derivative pattern versus local binary pattern : face recognition with high-order local pattern descriptor», *IEEE transactions on image processing*, vol. 19, n° 2, p. 533–544. [34](#), [35](#)
- ZHANG, J., M. MARSZALEK, S. LAZEBNIK et C. SCHMID. 2007, «Local features and kernels for classification of texture and object categories : A comprehensive study», *International journal of computer vision*, vol. 73, n° 2, p. 213–238. [23](#)
- ZHANG, W., S. SHAN, W. GAO, X. CHEN et H. ZHANG. 2005, «Local gabor binary pattern histogram sequence (lgbphs) : a novel non-statistical model for face representation and recognition», dans *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1, IEEE, p. 786–791. [32](#), [35](#)
- ZHAO, G., T. AHONEN, J. MATAS et M. PIETIKAINEN. 2012a, «Rotation-invariant image and video description with local binary pattern features», *IEEE Transactions on Image Processing*, vol. 21, n° 4, p. 1465–1477. [35](#), [36](#)
- ZHAO, Y. 2011, «Theories and applications of lbp : a survey», dans *International Conference on Intelligent Computing*, Springer, p. 112–120. [25](#)
- ZHAO, Y., D.-S. HUANG et W. JIA. 2012b, «Completed local binary count for rotation invariant texture classification», *IEEE transactions on image processing*, vol. 21, n° 10, p. 4492–4497. [32](#), [34](#), [35](#), [58](#), [83](#), [111](#), [118](#)
- ZHAO, Z., X. HUANG et G. YANG. 2015, «Plant recognition based on leaf and bark images», *Journal of Computational Information Systems*, vol. 11, n° 3, p. 857–864. [14](#), [15](#), [38](#)

ZHU, H., X. HUANG, S. ZHANG et P. C. YUEN. 2017, «Plant identification via multipath sparse coding», *Multimedia Tools and Applications*, vol. 76, n° 3, p. 4599–4615. [14](#), [15](#), [38](#)

Bibliographie de l'auteur

Publications internationales

- ▶ Boudra S., Yahiaoui I., Behloul A. (2021) A set of statistical radial binary patterns for tree species identification based on bark images. *Multimed Tools Appl* 80, 22373–22404. <https://doi.org/10.1007/s11042-020-08874-x>

Communications internationales

- ▶ Boudra S., Yahiaoui I., Behloul A. (2015) A Comparison of Multi-scale Local Binary Pattern Variants for Bark Image Retrieval. In : Battiato S., Blanc-Talon J., Gallo G., Philips W., Popescu D., Scheunders P. (eds) *Advanced Concepts for Intelligent Vision Systems. ACIVS 2015. Lecture Notes in Computer Science*, vol 9386. Springer, Cham. https://doi.org/10.1007/978-3-319-25903-1_66
- ▶ Boudra S., Yahiaoui I., Behloul A. (2017) Statistical Radial Binary Patterns (SRBP) for Bark Texture Identification. In : Blanc-Talon J., Penne R., Philips W., Popescu D., Scheunders P. (eds) *Advanced Concepts for Intelligent Vision Systems. ACIVS 2017. Lecture Notes in Computer Science*, vol 10617. Springer, Cham. https://doi.org/10.1007/978-3-319-70353-4_9
- ▶ Boudra S., Yahiaoui I., Behloul A. (2018) Bark Identification Using Improved Statistical Radial Binary Patterns. In *International Conference on Content-Based Multimedia Indexing CBMI 2018*, pp. 1-6, [doi:10.1109/CBMI.2018.8516536](https://doi.org/10.1109/CBMI.2018.8516536).
- ▶ Boudra S., Yahiaoui I., Behloul A. (2018) Plant identification from bark : A texture description based on Statistical Macro Binary Pattern," 2018 24th International Conference on Pattern Recognition (ICPR 2018), pp. 1530-1535, [doi:10.1109/ICPR.2018.8545798](https://doi.org/10.1109/ICPR.2018.8545798).