

CONTRIBUTION A LA REDUCTION DE LA DIMENSIONNALITE ET LA CLASSIFICATION DES IMAGES HYPERSPECTRALES EN UTILISANT L'INFORMATION MUTUELLE ET LA TEXTURE

Résumé : L'imagerie hyperspectrale de télédétection (IHS) permet d'acquérir des centaines de bandes pour la même région. Ainsi, un spectre complet de réflectance est construit pour chaque pixel de la scène. Cette grande quantité de données augmente la discrimination des objets à classifier. Cependant, deux défis s'imposent : la malédiction de dimensionnalité et le problème de pertinence des bandes nécessaires pour distinguer les classes. La problématique consiste à réduire la dimensionnalité des IHS pour augmenter la performance de la classification. Dans le cadre du travail de cette thèse, nous avons proposé des méthodes originales pour trouver le groupe réduit d'attributs « bandes » les plus informatifs et pertinents à la classification des IHS. D'une part, nous avons exploité l'information spectrale en utilisant l'information mutuelle normalisée et la probabilité d'erreur dans des approches filtres et Wrapper. Des seuils sont introduits pour le contrôle de la redondance utile. D'autre part, nous avons proposé l'utilisation conjointe de l'information spectrale-spatiale en ajoutant les caractéristiques de textures des bandes. Le classifieur SVM a été retenu pour le développement de nos approches proposées suite à l'évaluation de plusieurs classifieurs. Les approches introduites ont été validées en utilisant trois images hyperspectrales réelles fournies par le capteur hyperspectral AVIRIS de la NASA et le capteur ROSIS. Ces méthodes proposées améliorent la performance de classification avec un faible coût de calcul en comparaison avec les méthodes récentes.

Mots clés : Images hyperspectrales, classification, réduction de dimensionnalité, sélection d'attributs, approches filtre-wrapper, information mutuelle normalisée, texture.

Abstract: Hyperspectral remote sensing imagery (IHS) allows the acquisition of hundreds of bands for the same region. Thus, a complete reflectance spectrum is constructed for each pixel of the scene. This large amount of data increases the discrimination of the objects to be classified. However, two challenges are faced: the curse of dimensionality and the problem of the relevance of the bands needed to distinguish the classes. The problematic is to reduce the dimensionality of IHS in order to increase the classification performance. As part of the work of this thesis, we proposed original methods for finding the smallest group of attributes «bands» that are most informative and relevant to the classification of IHS. On the one hand, we have exploited spectral information by using normalized mutual information and error probability in filter and Wrapper approaches. Thresholds are introduced for controlling useful redundancy. On the other hand, we proposed the use of both spectral and spatial information by adding the texture characteristics of the bands. The SVM classifier was chosen for the development of our proposed approaches based on the evaluation of several classifiers. The introduced approaches were validated using three real benchmark hyperspectral images provided by the NASA's AVIRIS hyperspectral sensor and the ROSIS sensor. These proposed methods improve the classification performance with a reduced processing time compared to recent methods.

Keywords: Hyperspectral images, classification, dimensionality reduction, attribute selection, filter-wrapper approaches, normalized mutual information, texture.

Hasna NHAILA

Contribution à la réduction de la dimensionnalité et la classification des images hyperspectrales en utilisant l'information mutuelle et la texture

N° thèse : 145/ST2I

Année : 2019

Année : 2019



Thèse N° : 145/ST2I

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

THÈSE DE DOCTORAT

CONTRIBUTION A LA REDUCTION DE LA DIMENSIONNALITE
ET LA CLASSIFICATION DES IMAGES HYPERSPECTRALES
EN UTILISANT L'INFORMATION MUTUELLE ET LA TEXTURE

Présentée par

Hasna NHAILA

Le 06/07/2019

Formation doctorale : Sciences de l'Ingénieur – Génie Electrique
Structure de recherche : Laboratoire de Recherche en Génie Electrique,
Equipe : Electronic Systems, Sensors and Nanobiotechnologies (E2SN),
ENSET de Rabat

JURY

Professeur Larbi BELLARBI

PES, ENSET, Université Mohammed V de Rabat

Président

Professeur Ahmed HAMMOUCH

PES, ENSET, Université Mohammed V de Rabat

Directeur de thèse

Professeur Raja TOUAHNI

PES, FS, Université Ibn Tofail, Kénitra

Rapporteur

Professeur Moulay Driss RAHMANI

PES, FS, Université Mohammed V de Rabat

Rapporteur

Professeur Jamal EL MHAMDI

PES, ENSET, Université Mohammed V de Rabat

Examineur

*« On enseigne les métiers, les sciences, les arts,
mais les mœurs et la vertu s'inspirent »*

-Citation de Marie

Joseph CERUËR

Dédicaces

C'est avec fierté et gratitude que j'entame ma thèse avec cette dédicace

A mes très chers parents

Tous les mots ne sauraient exprimer mon amour, ma gratitude et ma reconnaissance pour vos efforts, vos sacrifices et vos prières grâce auxquelles j'ai pu terminer ce travail dans l'enthousiasme. Que Dieu tout puissant vous procure santé, longue vie et bonheur éternel.

A mon mari

Aucune dédicace ne pourrait traduire l'ampleur de ma reconnaissance vis-à-vis de ton affection, ton soutien et ta compréhension. Je te dédie ce travail en guise de mon respect et mon admiration, que Dieu te procure longue vie et santé.

A ma princesse Rim

Tu es l'éclat de ma vie, je t'aime ma fille. Que Dieu tout puissant te protège et te donne longue vie, pour que je puisse te voir grandir et parcourir ta vie dans la joie et le bonheur.

A mon frère et mes sœurs

Je vous dédie ce travail en témoignage de votre implication, votre amour et votre soutien moral. J'ai de la chance de vous avoir à mes côtés,

A mes ami(e)s et tous ceux qui me sont chers

Je vous souhaite à travers ce travail une vie pleine de santé, de joie et de réussite

A mes enseignants tout au long de mes études et à tous ceux que ma réussite leur tient à cœur.

A vous tous je dédie ce travail

Remerciements

Mes sincères remerciements s'adressent en premier lieu à mon cher professeur Ahmed HAMMOUCH, mon directeur de thèse qui m'a transmis la passion de la recherche. Je vous remercie pour votre disponibilité, vos remarques pertinentes, et vos conseils fructueux tout au long de ma période de recherche.

Ma profonde gratitude s'adresse ensuite au Docteur Elkebir SARHROUNI pour son aide précieuse, ses critiques bienveillantes et son suivi perspicace ainsi que son soutien scientifique et moral durant ces années de travail.

Je tiens à remercier également tous les membres du jury pour l'honneur qu'ils m'ont accordé en acceptant de juger ce travail. Qu'ils trouvent ici l'expression de ma forte reconnaissance.

Mes chaleureux remerciements s'adressent aussi à tout le corps professoral et administratif de l'ENSET. Nous avons trouvé auprès de vous une grande compétence, une orientation avisée et une conscience professionnelle. Que vous trouvez ici notre profond respect.

Je ne saurais terminer sans remercier vivement tous mes collègues. Qu'ils trouvent ici toute ma reconnaissance pour leurs échanges constructifs et leur amitié.

Enfin, mon respectueux dévouement s'adresse à toute personne qui a contribué, de près ou de loin, à la réalisation de ce mémoire.

Résumé

L'imagerie hyperspectrale de télédétection (IHS) permet d'acquérir des centaines de bandes pour la même région. Ainsi, un spectre complet de réflectance est construit pour chaque pixel de la scène. Cette grande quantité de données augmente la discrimination des objets à classifier. Cependant, deux défis s'imposent : la malédiction de dimensionnalité et le problème de pertinence des bandes nécessaires pour distinguer les classes. La problématique consiste à réduire la dimensionnalité des IHS pour augmenter la performance de la classification. Dans le cadre du travail de cette thèse, nous avons proposé des méthodes originales pour trouver le groupe réduit d'attributs « bandes » les plus informatifs et pertinents à la classification des IHS. D'une part, nous avons exploité l'information spectrale en utilisant l'information mutuelle normalisée et la probabilité d'erreur dans des approches filtres et Wrapper. Des seuils sont introduits pour le contrôle de la redondance utile. D'autre part, nous avons proposé l'utilisation conjointe de l'information spectrale-spatiale en ajoutant les caractéristiques de textures des bandes. Le classifieur SVM a été retenu pour le développement de nos approches proposées suite à l'évaluation de plusieurs classifieurs. Les approches introduites ont été validées en utilisant trois images hyperspectrales réelles fournies par le capteur hyperspectral AVIRIS de la NASA et le capteur ROSIS. Ces méthodes proposées améliorent la performance de classification avec un faible coût de calcul en comparaison avec les méthodes récentes.

Mots clés : Images hyperspectrales, classification, réduction de dimensionnalité, sélection d'attributs, approches filtre-wrapper, information mutuelle normalisée, texture.

Abstract

Hyperspectral remote sensing imagery (IHS) allows the acquisition of hundreds of bands for the same region. Thus, a complete reflectance spectrum is constructed for each pixel of the scene. This large amount of data increases the discrimination of the objects to be classified. However, two challenges are faced: the curse of dimensionality and the problem of the relevance of the bands needed to distinguish the classes. The problematic is to reduce the dimensionality of IHS in order to increase the classification performance. As part of the work of this thesis, we proposed original methods for finding the smallest group of attributes «bands» that are most informative and relevant to the classification of IHS. On the one hand, we have exploited spectral information by using normalized mutual information and error probability in filter and Wrapper approaches. Thresholds are introduced for controlling useful redundancy. On the other hand, we proposed the use of both spectral and spatial information by adding the texture characteristics of the bands. The SVM classifier was chosen for the development of our proposed approaches based on the evaluation of several classifiers. The introduced approaches were validated using three real benchmark hyperspectral images provided by the NASA's AVIRIS hyperspectral sensor and the ROSIS sensor. These proposed methods improve the classification performance with a reduced processing time compared to recent methods.

Keywords: Hyperspectral images, classification, dimensionality reduction, attribute selection, filter-wrapper approaches, normalized mutual information, texture.

Table des matières

Dédicaces	i
Remerciements	ii
Résumé	iii
Abstract	iv
Liste des figures.....	ix
Liste des tableaux	xi
Liste des abréviations	xiii
Introduction générale.....	1
Contexte général.....	1
Problématique.....	1
Objectif principal.....	2
Principales contributions	2
Organisation de la thèse	4
Chapitre1: Concept et problématiques liées aux images hyperspectrales.	6
1.1 Introduction	7
1.2 Concept de l'imagerie hyperspectrale	7
1.2.1 Principe de la télédétection.....	7
1.2.2 Caractéristiques des images hyperspectrales.....	9
1.2.3 Principales applications des IHS	10
1.3 Acquisition et prétraitement des IHS	11
1.3.1 Acquisition des IHS.....	11
1.3.2 Prétraitement des IHS.....	13
1.4 Problématiques relatives aux images hyperspectrales.....	14
1.5 Les images hyperspectrales d'application	15
1.5.1 AVIRIS_Indian Pines.....	16
1.5.2 AVIRIS_Salinas	17
1.5.3 ROSIS_Pavia University	18
1.6 Conclusion.....	19
Chapitre 2: État de l'art des méthodes de réduction de dimensionnalité et la classification des images hyperspectrales	20
2.1 Introduction	21
2.2 Les approches de classification des images hyperspectrales.....	21
2.2.1 Classification selon l'utilisation des échantillons d'entraînement	21

2.2.1.1 Méthodes supervisées.....	22
2.2.1.2 Méthodes non supervisées.....	23
2.2.2 Classification basée sur des paramètres.....	24
2.2.2.1 Méthodes paramétriques.....	24
2.2.2.2 Méthodes non paramétriques.....	25
2.2.3 Classification basée sur l'information des pixels.....	25
2.2.3.1 Méthodes par pixel 'Per-pixel'.....	25
2.2.3.2 Méthodes Sous-pixel 'Subpixel'.....	26
2.2.3.3 Méthodes Par-champ 'Per-field'.....	26
2.3 Les approches de réduction de dimensionnalité.....	27
2.3.1 Pourquoi la réduction de dimensionnalité?.....	27
2.3.2 Approches de réduction selon la génération d'attributs.....	28
2.3.2.1 Méthodes basées sur la sélection d'attributs.....	28
2.3.2.2 Méthodes basées sur l'extraction d'attributs.....	29
2.3.3 Approches de réduction selon la dépendance au classifieur.....	29
2.3.3.1 Approches Filtres.....	30
2.3.3.2 Approches Wrapper.....	30
2.3.3.3 Approches Hybrides.....	31
2.4 Critiques et limitations.....	31
2.5 Conclusion.....	32
Chapitre 3: Évaluation des performances des classifieurs : SVM, RF, KNN et LDA pour la classification des images hyperspectrales en utilisant l'information mutuelle IM.....	34
3.1 Introduction.....	35
3.2 Les recherches antérieures.....	35
3.3 Méthodologie proposée.....	37
3.3.1 Principe.....	37
3.3.2 Réduction de dimensionnalité des IHS par l'information mutuelle.....	38
3.3.2.1 Mesure de l'information mutuelle.....	38
3.3.2.2 Présentation de la sélection des bandes spectrales basée sur l'IM.....	39
3.3.3 Classifieurs étudiés.....	40
3.3.3.1 Machines à Vecteurs de Support SVM.....	40
3.3.3.2 K- plus proches Voisins K-NN « K-Nearests Neighbors ».....	42
3.3.3.3 Analyse discriminante linéaire LDA.....	43
3.3.3.4 Les forêts aléatoires « Random Forest RF ».....	44
3.4 Résultats et discussion.....	45

3.4.1 Description des bases de données hyperspectrales utilisées.....	45
3.4.2 Paramètres de classification et d'évaluation	46
3.4.3 Résultats et analyses	47
3.4.3.1 Comparaison des taux de classification des algorithmes selon le nombre de bandes sélectionnées.....	47
3.4.3.2 Comparaison des algorithmes selon les métriques : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées.....	52
3.4.3.3 Visualisation des scènes classifiées.....	55
3.5 Conclusion.....	57
Chapitre 4: Une nouvelle approche filtre pour la sélection de bandes et la classification des images hyperspectrales basée sur l'information mutuelle normalisée et les machines à vecteurs de support... 58	
4.1 Introduction	59
4.2 Les recherches antérieures.....	59
4.3 Méthodologie proposée	60
4.3.1 Principe.....	60
4.3.2 Sélection des bandes en utilisant l'information mutuelle normalisée	61
4.3.2.1 Formulation de l'information mutuelle normalisée NMI	61
4.3.2.2 Sélection des bandes pertinentes	62
4.3.3 Algorithme proposé.....	64
4.4 Résultats expérimentaux et discussion	66
4.4.1 Bases de données utilisées.....	66
4.4.2 Protocole expérimental.....	67
4.4.3 Résultats et discussion.....	67
4.5 Conclusion.....	71
Chapitre 5: Nouvelle méthode Wrapper basée sur l'information mutuelle normalisée pour la réduction de dimensionnalité et la classification des images hyperspectrales 72	
5.1 Introduction	73
5.2 Les recherches antérieures.....	73
5.3 Méthodologie proposée	74
5.3.1 Principe.....	74
5.3.2 Probabilité d'erreur PE.....	75
5.3.3 Application aux IHS	76
5.3.4 L'algorithme proposé	78
5.4 Expériences	79
5.4.1 Protocole expérimental.....	79
5.4.2 Résultats et discussion.....	79

5.5 Conclusion.....	85
Chapitre 6: Apport de la texture pour la réduction de dimensionnalité et la classification des images hyperspectrales : une nouvelle méthode basée sur les caractéristiques GLCM et l'information mutuelle	86
6.1 Introduction	87
6.2 Les méthodes de classification basées sur la texture	87
6.3 Méthodologie proposée	89
6.3.1 Principe.....	89
6.3.2 Sélection des bandes par IM et caractéristiques de texture	90
6.3.2.1 Rappel sur l'information mutuelle.....	90
6.3.2.2 Extraction des caractéristiques de texture	90
6.3.2.2.1 Contraste.....	91
6.3.2.2.2 Corrélation.....	92
6.3.2.2.3 Énergie.....	94
6.3.2.2.4 Homogénéité	95
6.3.3 L'algorithme proposé	97
6.4 Résultats expérimentaux.....	98
6.4.1 Protocol expérimental.....	98
6.4.2 Résultats et discussion.....	98
6.4.2.1 Résultats et analyse de comparaison des algorithmes MIBS et GLMI selon le nombre de bandes sélectionnées et différents seuils de redondance	99
6.4.2.1.1 Tableaux et figures des résultats.....	99
6.4.2.1.2 Analyse des résultats	106
6.4.2.2 Résultats et analyse de la classification des classes des IHS selon les critères de GLMI : Contraste, Corrélation, Énergie et Homogénéité	107
6.4.2.2.1 Résultats et analyse pour l'IHS Indian-Pines	107
6.4.2.2.2 Résultats et analyse pour l'IHS Salinas	109
6.4.2.2.3 Résultats et analyse pour l'IHS Pavia-University	111
6.4.2.3 Résultats et analyse de la comparaison de l'algorithme proposé GLMI avec les algorithmes MIM, MIBS et DSEBS pour 64 bandes.	113
6.5 Conclusion.....	115
Conclusion générale	116
Publications Scientifiques	118
Annexe	119
Références	122

Liste des figures

Figure 1. 1 : <i>La télédétection (a) passive, (b) active</i>	7
Figure 1. 2 : <i>Le spectre électromagnétique</i>	8
Figure 1. 3: <i>Principe de l'imagerie hyperspectrale</i>	10
Figure 1. 4: <i>Les modes d'acquisition du cube hyperspectral : (a) balayage de point ou whiskbroom, (b) balayage de ligne ou pushbroom et (c) balayage de zone (tunable filters), modifié d'après [Adão, 2017].</i>	11
Figure 1. 5: <i>L'imagerie hyperspectrale et les effets atmosphériques.</i>	14
Figure 1. 6: <i>Illustration du phénomène de Hughes [Hughes, 1968].</i>	15
Figure 1. 7: <i>Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale AVIRIS_Indian Pines.</i>	16
Figure 1. 8: <i>Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale AVIRIS_Salinas.</i>	17
Figure 1. 9: <i>Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale ROSIS_Pavia University.</i>	19
Figure 2. 1: <i>Les approches de classification des IHS</i>	22
Figure 2. 2: <i>Les approches de réduction de dimensionnalité des IHS</i>	28
Figure 2. 3: <i>Principe du modèle Filtre</i>	30
Figure 2. 4: <i>Principe du modèle Wrapper</i>	31
Figure 3. 1: <i>Schéma de principe de la méthodologie.</i>	37
Figure 3. 2: <i>Diagramme de Venn de la relation entre les entropies et l'IM de X et Y.</i>	38
Figure 3. 3: <i>Exemples de bandes d'AVIRIS_Indian Pines avec la vérité de terrain correspondante.</i> ..	45
Figure 3. 4: <i>Exemples de bandes d'AVIRIS_Salinas avec la vérité de terrain correspondante.</i>	46
Figure 3. 5: <i>Exemples de bandes de ROSIS_Pavia-University avec la vérité de terrain correspondante.</i>	46
Figure 3. 6: <i>Comparaison des performances en termes du taux de classification et des bandes sélectionnées pour l'image Indian-Pines</i>	51
Figure 3. 7: <i>Comparaison des performances en termes de précision de la classification et des bandes sélectionnées pour la scène Salinas.</i>	51
Figure 3. 8: <i>Comparaison des performances en termes de précision de la classification et des bandes sélectionnées pour Pavia-University.</i>	51
Figure 3. 9: <i>La vérité de terrain de Indian-Pines (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.</i>	56
Figure 3. 10: <i>La vérité de terrain de Salinas (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.</i>	56
Figure 3. 11: <i>La vérité de terrain de Pavia (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.</i>	56
Figure 4. 1: <i>Organigramme de la méthode de sélection avec validation proposée</i>	60
Figure 4. 2: <i>L'Information mutuelle normalisée NMI et l'Information mutuelle IM des bandes spectrales d'Indian-Pines (à gauche) et Salinas (à droite), avec les GT correspondantes</i>	63

Figure 4. 3: L'image couleur composite et la vérité de terrain correspondante avec les classes pour: les IHS (a) Indian Pines et (b) Salinas.	66
Figure 4. 4: La vérité-terrain d'Indian-Pines (a) et les images classifiées en utilisant la méthode proposée pour différents pourcentages de données d'entraînement: 50% (b), 25% (c) et 10% en (d). 69	69
Figure 4. 5: La vérité-terrain de Salinas (a) et les images classifiées en utilisant la méthode proposée pour différents pourcentages de données d'entraînement: 50% (b), 25% (c) et 10% en (d).	69
Figure 5. 1: Organigramme de la méthode Wrapper proposée.....	74
Figure 5. 2: La probabilité d'erreur calculée avec l'IM pour chaque bande avec GT	77
Figure 5. 3: L'intervalle PE* calculé avec la NMI pour chaque bande avec GT	77
Figure 5. 4: L'image hyperspectrale Indian-Pines: (a) Carte de la vérité de terrain ; et les cartes classifiées obtenues pour 60 bandes sélectionnées en utilisant les algorithmes (b) MRMR (OA=83.95%), (c) NMIBS (88.63%) et les méthodes proposées WNMPIE en (d) avec (OA=89.17%) et WNMPIE* (e) avec (OA=93.34%).	83
Figure 5. 5 : L'image hyperspectrale Salinas: (a) Carte de la vérité de terrain ; et les cartes classifiées obtenues pour 37 bandes sélectionnées avec une précision globale OA (entre parenthèses) en utilisant les algorithmes reproduits (b) MRMR (88.71%), (c) NMIBS (91.41%), (d) WNMPIE (91.63%) et (e) WNMPIE* (95.56).	84
Figure 6. 1: Méthodes de classification de texture	88
Figure 6. 2: L'organigramme de la méthodologie proposée	89
Figure 6. 3: Image 5x5 avec 3 niveaux de gris et la matrice de cooccurrence correspondante	90
Figure 6. 4: Le contraste et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, avec trois bandes et leurs valeurs correspondantes de Contraste et d'IM	91
Figure 6. 5 : La corrélation et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, avec trois bandes et leurs valeurs correspondantes de corrélation et d'IM.....	93
Figure 6. 6: L'énergie et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'énergie et d'IM.	94
Figure 6. 7: L'homogénéité et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'homogénéité et d'IM.....	96
Figure 6. 8 : Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur AVIRIS Indian-Pines.	101
Figure 6. 9 : Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur AVIRIS Salinas.	103
Figure 6. 10: Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur ROSIS_Pavia University	105
Figure 6. 11: La Vérité-terrain d'Indian-Pines (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D).	108
Figure 6. 12: La Vérité-terrain de Salinas (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D).....	110
Figure 6. 13: La Vérité-terrain de Pavia (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D).	112
Figure 6. 14: Graphiques comparatifs des performances des trois images Indian, Salinas et Pavia en utilisant les différentes méthodes en termes d'OA et de Kappa.	114

Liste des tableaux

Tableau 1. 1: <i>Caractéristiques techniques des capteurs hyperspectraux AVIRIS et ROSIS</i>	13
Tableau 1. 2: <i>Nom et Nombre de pixels de chaque classe pour l'image AVIRIS_Indian Pines</i>	17
Tableau 1. 3: <i>Nom et Nombre de pixels de chaque classe pour l'image AVIRIS_Salinas</i>	18
Tableau 1. 4: <i>Nom et nombre de pixels de chaque classe pour l'image ROSIS_Pavia University</i>	19
Tableau 3. 1: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas d'Indian-Pines.</i>	48
Tableau 3. 2: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas de Salinas</i>	49
Tableau 3. 3: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas de Pavia-University</i>	50
Tableau 3. 4: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas d'Indian-Pines.</i>	53
Tableau 3. 5: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas de Salinas</i>	53
Tableau 3. 6: <i>Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas de Pavia-University.</i>	54
Tableau 4. 1: <i>Précision globale de classification OA et précision moyenne de classification AA selon le groupe de bandes sélectionnées en utilisant l'IM et la NMI pour Indian-Pines.</i>	64
Tableau 4. 2: <i>Caractéristiques des images hyperspectrales utilisées dans les expériences.</i>	66
Tableau 4. 3: <i>La précision globale OA Obtenue en utilisant l'algorithme proposé sur l'image hyperspectrale Indian-Pines pour différents nombre de bandes sélectionnées et différents pourcentages d'échantillons d'apprentissage.</i>	68
Tableau 4. 4: <i>La précision globale OA Obtenue en utilisant l'algorithme proposé sur l'image hyperspectrale Salinas pour différents nombre de bandes sélectionnées et de pourcentage d'échantillons d'apprentissage.</i>	68
Tableau 4. 5: <i>La précision moyenne AA (%), la précision globale OA (%) et le temps de calcul (s) obtenus par l'algorithme proposé en comparaison avec différentes méthodes sur les bases de données Indian-Pines pour 75 bandes sélectionnées et Salinas pour 69 bandes sélectionnées.</i>	70
Tableau 5. 1: <i>AA(%), OA(%) et le coefficient kappa obtenus par la méthode proposée sur l'image Indian-Pines en utilisant 49 bandes sélectionnées avec trois cas d'ensembles d'apprentissage.</i>	80
Tableau 5. 2: <i>AA(%), OA(%) et le coefficient kappa obtenus par la méthode proposée sur l'image Salinas en utilisant 37 bandes sélectionnées avec trois cas d'ensembles d'apprentissage.</i>	80
Tableau 5. 3: <i>Les performances de la classification de l'IHS Indian-Pines en utilisant la méthode proposée en comparaison avec les algorithmes MRMR et NMIBS pour 60 bandes sélectionnées.</i>	82
Tableau 5. 4: <i>Les performances de la classification de l'IHS Salinas en utilisant la méthode proposée en comparaison avec les algorithmes MRMR et NMIBS pour 37 bandes sélectionnées.</i>	82

Tableau 6. 1: <i>Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Indian-Pines.</i>	100
Tableau 6. 2: <i>Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Salinas.</i>	102
Tableau 6. 3: <i>Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Pavia-University.</i>	104
Tableau 6. 4: <i>La précision de la classification individuelle ICA (%) de chaque classe de l'image Indian-Pines en utilisant l'algorithme proposé pour $Th=-0,02$</i>	108
Tableau 6. 5: <i>La précision de la classification individuelle ICA (%) de chaque classe de l'image Salinas en utilisant l'algorithme proposé pour $Th=-0,02$</i>	110
Tableau 6. 6: <i>La précision de la classification individuelle ICA (%) de chaque classe de l'image Pavia en utilisant l'algorithme proposé pour $Th=-0,02$</i>	111
Tableau 6. 7: <i>La précision globale OA (%), le coefficient Kappa k et le temps d'exécution $T(s)$ des trois images obtenus par les différentes méthodes en utilisant 64 bandes sélectionnées.</i>	113

Liste des abréviations

- AA Précision moyenne de classification « Average Accuracy »
- GA Algorithmes Génétiques
- GLCM Matrice de Co-Occurrence des Niveaux de Gris « Gray Level Co-Occurrence Matrix »
- GT Vérité de terrain « Ground Truth »
- ICA Précision Individuelle de classification « Individual class Accuracy »
- IHS Images Hyperspectrales
- IM Information Mutuelle
- KNN K Plus Proches Voisins
- LDA Analyse Discriminante Linéaire
- NMI Information Mutuelle Normalisée
- OA Précision globale de classification « Overall Accuracy »
- PBS Sélection progressive de bandes
- PE Probabilité d'erreur
- RBF Fonction de Base Radiale «Radial Basis Function »
- RF Forêt Aléatoire « Random Forest »
- SVM Machines à Vecteurs de Support

Introduction générale

Contexte général

La télédétection par imagerie hyperspectrale est une technologie en plein développement qui permet l'acquisition des informations spectrales et spatiales d'un objet cible. Le système fournit plusieurs centaines de bandes (images) étroites et contiguës de la même scène observée. Ces bandes couvrent la région du visible aux ondes courtes infrarouges du spectre électromagnétique (400 - 2500 nm). Elles sont regroupées sous forme d'un hypercube 3D, dont deux dimensions spatiales et une dimension spectrale selon les différentes longueurs d'onde.

Avec le développement récent des technologies des capteurs hyperspectraux, les images hyperspectrales (IHS) sont devenues plus disponibles et trouvent ainsi leurs applications dans différents domaines ; tels que l'industrie alimentaire [Dai, 2015 ; Feng, 2018], la médecine [Lu, 2014 ; Danielis, 2017], l'astronomie [Hege, 2004], la cartographie de l'agriculture et plus particulièrement l'analyse de la couverture terrestre [Frohn, 2017 ; Laurin, 2016 ; Schmidt, 2017].

Dans les applications du monde réel en télédétection, la grande quantité d'informations contenue dans les IHS augmente la discrimination et la classification des différents objets observés de la terre. Elle permet également une analyse fine des caractéristiques physiques et chimiques des zones étudiées [ElMasry, 2016].

Dans cette thèse, nous nous intéresserons à la production des cartes thématiques des surfaces terrestres en guise d'observation et d'exploration des ressources de la terre.

Problématique

Dans les schémas de classification, la richesse des images hyperspectrales en termes de données améliore la caractérisation des différentes classes de la région observée. Cependant, de nombreux défis sont confrontés au niveau de la pertinence des données et du temps de traitement. Ceci est lié principalement à la forte corrélation entre les bandes adjacentes, la présence de bandes non pertinentes et du nombre fini des échantillons d'apprentissage. Ce problème est connu sous l'appellation : la malédiction de dimensionnalité ou encore phénomène de Hughes [Hughes, 1968]. En effet, le spectre des IHS contient des zones non informatives qui faussent l'apprentissage. D'autre part, des zones spectrales différentes

peuvent contenir une information redondante qui complique le processus d'apprentissage. La sélection des bandes pertinentes et non redondantes est une tâche primordiale pour une bonne classification des images hyperspectrales. Le test de toutes les combinaisons possibles pour trouver l'ensemble de bandes adéquates est pratiquement impossible. Dans ce cadre, différentes approches sont nécessaires pour trouver un ensemble adéquat de bandes pertinentes pour augmenter les performances de la classification.

Objectif principal

Dans le cadre de ces travaux de recherche, nous nous sommes intéressés aux approches utilisant l'information mutuelle (IM) et la texture dans des stratégies « Filtre » et « Wrapper » pour la sélection des bandes pertinentes. L'objectif est de proposer des méthodes originales, selon diverses approches, pour améliorer la réduction de dimensionnalité et la classification des images hyperspectrales de télédétection. Ces méthodes visent à reproduire des cartes thématiques. Le classifieur Machines à Vecteurs de Support (SVM) est choisi en se basant sur une étude comparative par expériences pour valider les algorithmes proposés. L'application est effectuée sur trois ensembles de données hyperspectrales fournies par le capteur (AVIRIS) de la NASA « Airborne Visible/Infrared Imaging Spectrometer Sensor » et le capteur (ROSIS) « Reflective Optics System Imaging Spectrometer ». Ces IHS ont des caractéristiques différentes en termes de types de données et de nombre de bandes et de classes. Elles sont largement utilisées comme référence pour tester la performance des algorithmes de classification des IHS dans de nombreuses recherches [Kang, 2015; Li 2017].

Principales contributions

Notre travail apporte cinq principales contributions portant sur la réduction de dimensionnalité et la classification des images hyperspectrales. Leurs résultats ont été publiés dans des conférences et journaux internationaux.

La première contribution consiste à introduire le concept de l'imagerie hyperspectrale, ses particularités, ses modes d'acquisition, son traitement ainsi que ses principales applications. Ceci nous permet de cerner les problématiques liées au traitement de ce type de données de haute dimensionnalité [Nhaila, 2014].

La deuxième contribution repose sur l'évaluation expérimentale des performances de quatre classifieurs en utilisant l'information mutuelle IM dans l'étape de réduction de dimensionnalité des IHS. Il s'agit des Machines à Vecteurs de Support SVM, les K plus

proches voisins (KNN), les forêts aléatoires « Random Forest » (RF) et l'Analyse Discriminante Linéaire (LDA). Ces algorithmes appartiennent à plusieurs catégories de méthodes de classification. Ils sont implémentés en tenant compte du caractère multi-classes des IHS. L'étude comparative est faite selon les critères d'évaluation les plus utilisés dans le domaine de la télédétection. Pour effectuer nos expériences, trois images hyperspectrales réelles ont été utilisées ; elles ont des caractéristiques différentes et sont fournies par deux capteurs hyperspectraux. Cette étude nous a permis de choisir le classifieur SVM pour la validation de nouvelles méthodes proposées dans ce mémoire, vu le compromis réalisé entre les performances de classification et le coût de calcul [Nhaila, 2018a].

La troisième contribution est une approche « Filtre » basée sur l'utilisation de l'information mutuelle normalisée (NMI) pour la sélection des bandes pertinentes [Nhaila, 2018b]. Le critère de choix stipule que la bande doit être retenue lorsque son NMI avec la vérité de terrain est élevée et qu'elle permet d'augmenter la NMI de la vérité estimée avec la vérité de terrain originale. L'ensemble des bandes choisies est soumis au classifieur SVM pour la validation et la reproduction de la carte thématique. L'algorithme est testé sur deux IHS réelles du capteur AVIRIS de la NASA. Les résultats obtenus sont satisfaisants. Elles ont permis de classifier avec succès les classes de la scène dans un intervalle réduit du temps d'exécution.

La quatrième contribution est une approche Wrapper basée sur l'information mutuelle normalisée NMI, le classifieur SVM et la probabilité d'erreur (PE) [Nhaila, 2018c]. Contrairement à la méthode proposée au chapitre précédent où la sélection de bandes se fait indépendamment de l'algorithme d'induction, cette approche intègre les performances du classifieur dans le processus de sélection de bandes pertinentes. Ceci est réalisé par le contrôle de la probabilité d'erreur calculée entre la vérité de terrain originale et celle reproduite par le SVM en utilisant l'ensemble de bandes sélectionnées. Ainsi, une bande est retenue si elle augmente l'information mutuelle normalisée et minimise l'intervalle de la probabilité d'erreur. La validation de la méthode proposée est faite sur deux images hyperspectrales du capteur AVIRIS. Les résultats obtenus sont très satisfaisants. Cependant, cet algorithme est coûteux, en termes du temps de calcul, vu qu'il fait appel plusieurs fois au classifieur SVM durant la sélection. Cette méthode est bénéfique pour les applications qui exigent une précision élevée de classification.

La cinquième contribution concerne l'apport de l'information spatiale aux performances de la classification des images hyperspectrales [Nhaila, 2015 ; Nhaila, 2018d]. En effet, les approches proposées dans les chapitres précédents utilisent l'information spectrale des bandes pour effectuer la sélection. Cette méthode en contrepartie, est basée sur l'utilisation conjointe de l'information spectrale-spatiale des bandes. Le processus de sélection de cet algorithme associe l'information mutuelle et les caractéristiques de texture des bandes. Nous analysons la texture à travers quatre caractéristiques (le contraste, la corrélation, l'énergie et l'homogénéité) pour améliorer la pertinence de l'ensemble des bandes sélectionnées. La validation est faite par le classifieur SVM sur trois images hyperspectrales fournies par le capteur hyperspectral AVIRIS de la NASA et le capteur hyperspectral ROSIS. Les résultats obtenus sont très prometteurs, ils ouvrent un axe de recherche dédié à l'analyse des caractéristiques de texture des bandes pour la réduction et la classification des images hyperspectrales.

Organisation de la thèse

Ce mémoire de thèse est composé de six chapitres suivis d'une conclusion générale.

Dans le premier chapitre, nous donnons le concept de l'imagerie hyperspectrale, ses caractéristiques, ses modes d'acquisition, son traitement ainsi que ses principales applications. Ceci va nous permettre de ressortir les problématiques relatives à la haute dimensionnalité des données issues des images hyperspectrales IHS.

Dans le deuxième chapitre, nous nous intéresserons à la problématique de réduction de dimensionnalité et la classification des IHS. Nous dressons un état de l'art des diverses méthodes développées par les chercheurs dans ce domaine. Enfin, nous retenons les approches qui seront adoptées dans ce mémoire pour le développement de nos propres algorithmes.

Le troisième chapitre présente une analyse expérimentale de différentes méthodes de classification supervisée. Le but est d'évaluer leurs performances aux problèmes de classification multi-classes relatives aux images hyperspectrales. L'information mutuelle IM est utilisée pour réduire la dimensionnalité des données d'entrée. Les résultats des applications sur trois IHS réelles de référence nous permettent un choix justifié du classifieur convenable à notre problématique.

Dans le quatrième chapitre, nous présentons une nouvelle méthode pour la réduction de dimensionnalité et la classification des IHS. Il s'agit d'une approche Filtre basée sur l'information mutuelle normalisée NMI et les machines à vecteurs de support SVM. Enfin, nous présentons l'avantage et la limitation de l'algorithme proposé.

Le cinquième chapitre propose une méthode originale pour pallier la limitation de la méthode précédente. Elle est basée sur une approche Wrapper qui utilise la NMI, le SVM et la probabilité d'erreur PE. Le but est d'améliorer la sélection des bandes pertinentes et diminuer la redondance dans les sous-ensembles sélectionnés. L'évaluation des performances de cette méthode est faite sur deux IHS du capteur AVIRIS.

Dans le sixième chapitre, nous introduisons une nouvelle approche Filtre qui combine l'information spectrale des bandes et l'information spatiale des pixels; les résultats obtenus sur trois IHS montrent que notre méthode surpasse les autres approches en augmentant les performances de la classification dans un temps réduit.

Enfin, nous clôturons notre mémoire par une conclusion générale qui synthétise les travaux réalisés et donne quelques perspectives de recherche à ces travaux.

A graphic element for the chapter title. It consists of a black rounded rectangle at the top containing the word "Chapitre" in white. Below this is a white rounded rectangle containing a large black number "1".

Chapitre
1

Concept et problématiques liées aux images hyperspectrales.

Résumé :

Dans ce premier chapitre, nous introduisons le concept de l'imagerie hyperspectrale (IHS): sa définition, ses caractéristiques, ses procédés d'acquisition, son traitement et ses principales applications. Cela nous permet de cerner les problématiques relatives à la haute dimensionnalité des données issues des IHS.

1.1 Introduction

Durant la dernière décennie, l'imagerie hyperspectrale IHS ou la spectro-imagerie a atteint des progrès notables en acquisition des données ayant une résolution spectrale plus élevée par rapport aux images multispectrales et monochromes. Cette nouvelle technologie nous a permis d'avoir un spectre continu pour chaque pixel de l'image. L'imagerie hyperspectrale trouve ses applications dans plusieurs domaines tels que la télédétection, l'imagerie médicale, l'agriculture, l'industrie des aliments, etc.

1.2 Concept de l'imagerie hyperspectrale

1.2.1 Principe de la télédétection

La télédétection est l'ensemble des techniques qui permettent l'observation et l'exploration de la terre au moyen d'instruments placés à distance. Cette technologie utilise la réflexion de la surface terrestre qui réagit différemment aux radiations qui proviennent des diverses sources lumineuses telles que le soleil dans le cas de la télédétection passive, ou des capteurs eux-mêmes dans le cas de la télédétection active, voir figure 1.1.

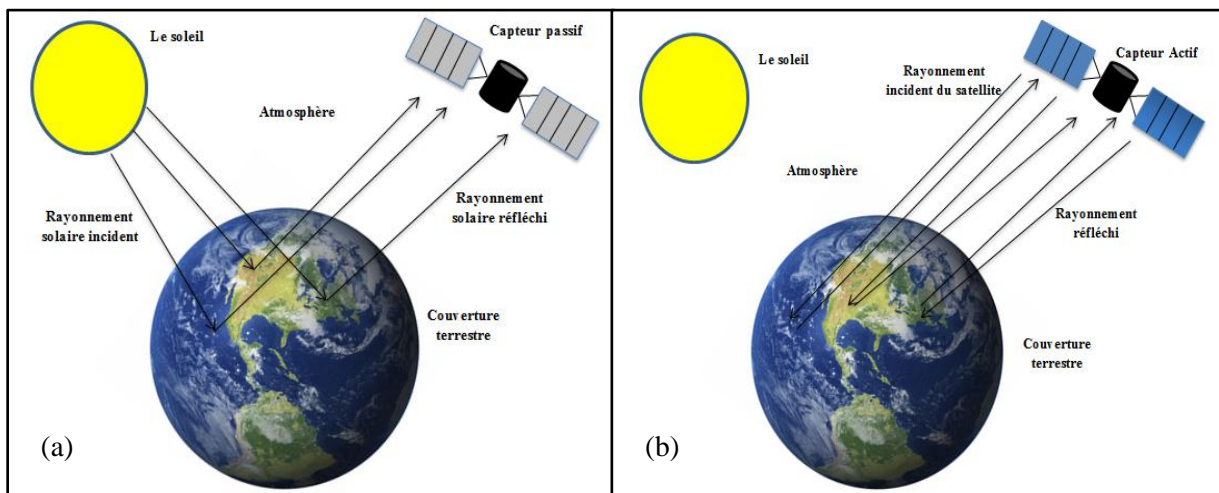


Figure 1. 1 : *La télédétection (a) passive, (b) active*

Pour l'identification et la caractérisation des objets cibles, on utilise des capteurs qui permettent de mesurer la lumière réfléchi. L'œil humain peut être considéré comme le premier capteur en télédétection qui permet de voir les lumières du champ visible entre le violet et le rouge (400nm à 700nm), tandis qu'une très grande partie du spectre électromagnétique, représenté en figure 1.2, reste invisible à l'œil nu, ce qui fait appel aux dispositifs numériques de télédétection pour pouvoir analyser les lumières au-delà du visible.

En effet, le rayonnement électromagnétique est caractérisé par sa longueur d'onde λ et sa fréquence f qui sont inversement proportionnelles, le spectre en figure 1.2 montre que les rayons sont répartis selon plusieurs régions :

- Rayons gamma et rayons X qui correspondent aux courtes longueurs d'onde <10 nm. Cette zone n'est pas disponible en télédétection vu que les rayons incidents sont totalement absorbés par la haute atmosphère.
- Rayons ultraviolets entre 10nm et 300nm qui sont aussi absorbés par l'ozone en atmosphère. Par contre, la région entre 300nm et 400nm dite photographique est détectable.
- Les lumières visibles à l'œil humain entre 400nm et 700nm.
- Rayons infrarouges (proche, moyen et thermique) entre 700nm et 1mm qui sont toutes accessibles en télédétection.
- Micro-ondes de 1mm à 1m, cette région est aussi importante pour les capteurs actifs que passifs.
- Les ondes radio au-delà de 1m jusqu'aux kilomètres qui nécessitent des radars spécifiques de très grande longueur d'onde.

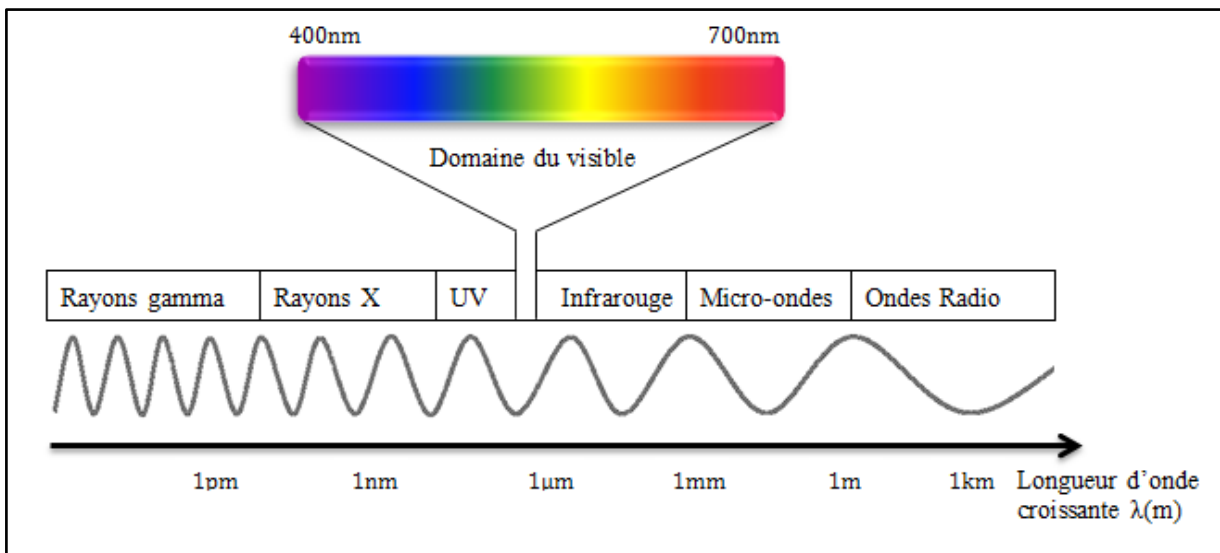


Figure 1. 2 : *Le spectre électromagnétique*

Plusieurs capteurs ont été développés en se basant sur le constat que l'observation des objets terrestres sur plusieurs parties du spectre précitées permet une meilleure identification des cibles. Les cameras couleur, par exemple, enregistrent les longueurs d'ondes dans la région de la lumière visible en se limitant sur trois bandes spectrales (Rouge : 650nm, Vert : 550nm et Bleu : 450nm) [Cordell, 2017]. Toutefois, en raison de similitude de couleur de quelques

objets de la terre, comme les types des végétations, les capteurs multispectraux qui enregistrent plus de bandes entre 3 et 10 deviennent plus efficaces que la technologie classique RVB. D'autre part, les capteurs hyperspectraux IHS constituent la révolution en télédétection, ils enregistrent les images en utilisant des centaines voire des milliers de bandes étroites du visible à l'infrarouge. La puissance de cette technologie réside dans l'amélioration de la compréhension des matériaux, tels que le sol, les roches et l'eau, etc. en fournissant des informations plus détaillées sur les zones étudiées.

1.2.2 Caractéristiques des images hyperspectrales

L'imagerie hyperspectrale est une technique émergente de télédétection. Elle permet d'acquérir simultanément des données spatiales et spectrales, dans plusieurs centaines de bandes étroites et contiguës, de la même région observée appelée vérité de terrain. Le domaine fréquentiel couvert par l'imagerie hyperspectrale débute du visible et s'étale vers la région des ondes infrarouges et proche infrarouge entre 400nm et 2500nm. Ces bandes sont combinées pour produire un cube des données à trois dimensions (spatiale-spectrale-spatiale) [Gowen, 2009]. La figure 1.3 représente un exemple d'hypercube constitué de deux dimensions spatiales dans le plan (X,Y) et la dimension spectrale selon les différentes longueurs d'onde λ . Ainsi, un spectre de réflectance quasi complet (nommé signature spectrale) est acquis à chaque pixel de la scène, ce qui permet d'analyser les caractéristiques physiques et chimiques pour distinguer les différents objets.

Les images hyperspectrales IHS se distinguent alors des images multispectrales et monochromes par trois principales caractéristiques [Chang, 2007]:

- Les IHS enregistrent des centaines de bandes spectrales de la même scène, tandis que les images multispectrales se limitent à un nombre de bandes entre trois et une dizaine.
- Les bandes acquises en utilisant les systèmes multispectraux sont irrégulièrement espacées à travers le spectre électromagnétique ; contrairement aux IHS qui fournissent des bandes régulièrement espacées et contiguës.
- La résolution spectrale (définie comme étant le rapport de la longueur d'onde centrale sur la largeur de la bande spectrale) des systèmes d'imagerie hyperspectrale est de l'ordre de 100, alors qu'elle est de l'ordre de 10 pour les images multispectrales. Cette caractéristique rend la discrimination des objets plus fine.

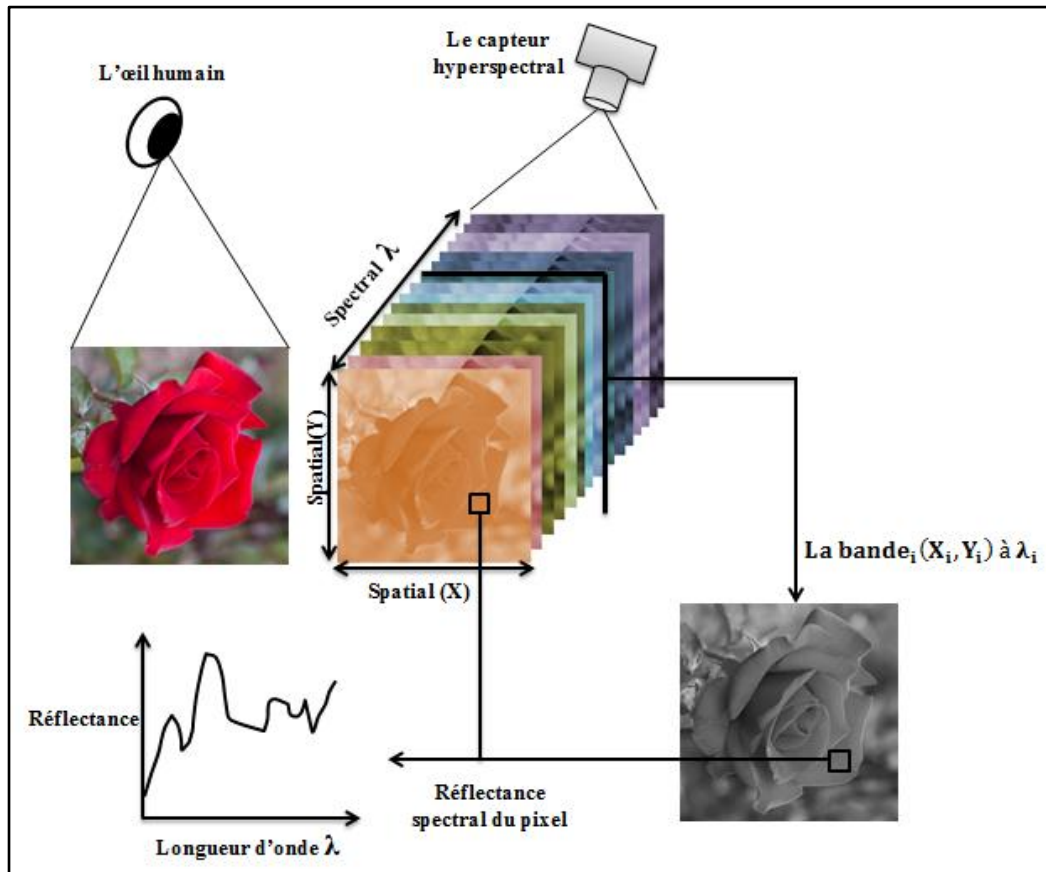


Figure 1. 3: Principe de l'imagerie hyperspectrale

1.2.3 Principales applications des IHS

L'imagerie hyperspectrale a été développée initialement pour la télédétection. Cependant, l'utilisation des IHS s'étend vers des domaines variés tels que l'astronomie et la surveillance de l'espace [Hege, 2004 ; Courbot, 2017], la pharmacologie [Dziewierz, 2016], la médecine pour le diagnostic des maladies [Lu, 2014 ; Danielis, 2017], la microbiologie et le traitement des aliments au cours de divers processus [Gowen, 2015 ; Liu, 2017] etc.

L'utilisation des IHS pour la télédétection nous fournit des données nécessaires à l'exploration des objets observés. Elle concerne des intérêts scientifiques et industriels [Nhaila, 2014]. Elle couvre aussi la surveillance des zones terrestres particulières telles que la détection des cibles et la cartographie des mines [Gonçalves, 2015 ; Makki, 2017]. Dans l'industrie des aliments, les images hyperspectrales représentent un moyen important pour l'analyse de l'authenticité, la qualité et la sécurité des produits alimentaires [Dai, 2015; ElMasry, 2016 ; Feng, 2018]. La détection des changements environnementaux et climatiques a aussi connu un avancement crucial via l'utilisation des IHS [Jay, 2017 ; Picón, 2017]. La classification et l'analyse de la couverture terrestre, qui représente notre domaine d'intérêt,

restent parmi les principales applications de l'imagerie hyperspectrale qui intéresse la recherche scientifique [Frohn, 2017 ; Laurin, 2016 ; Schmidt, 2017 ; Roscher, 2016].

1.3 Acquisition et prétraitement des IHS

1.3.1 Acquisition des IHS

Le système d'acquisition en télédétection hyperspectrale contient trois principales composantes : la source d'illumination, la plateforme et le capteur. La source principale qui illumine les scènes est le soleil : c'est une télédétection passive. Les capteurs sont placés sur des plateformes aériennes (avion, hélicoptère ...) ou spatiales. Pour chacun d'entre eux, la télédétection peut s'effectuer d'une manière active ou passive.

Les capteurs peuvent être classés en trois catégories qui donnent les modes d'acquisitions suivants: balayage de point, balayage de ligne et balayage de zone (longueur d'onde). Elles sont connues respectivement sous les appellations : « whiskbroom, pushbroom et tunable filters » [Liu, 2015]. Le principe de ces trois approches est illustré dans la figure 1.4.

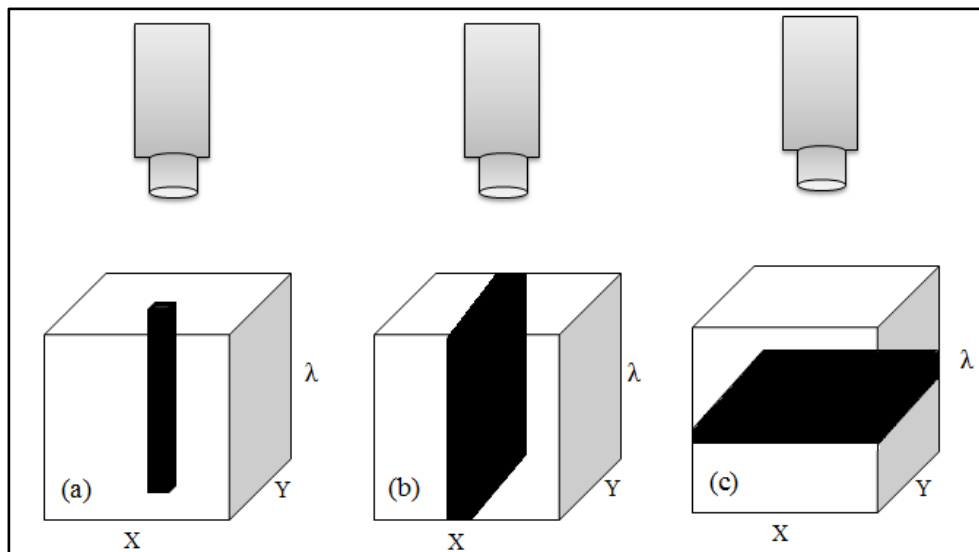


Figure 1. 4: Les modes d'acquisition du cube hyperspectral : (a) balayage de point ou whiskbroom, (b) balayage de ligne ou pushbroom et (c) balayage de zone (tunable filters), modifié d'après [Adão, 2017].

La technique « Whiskbroom » construit le cube hyperspectral de 3D en mesurant successivement le spectre complet de chaque pixel en se déplaçant selon deux dimensions spatiales ; les données sont stockées sous format « BIP ». Le capteur AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) utilise le mode balayage de point « whiskbroom ».

- L'approche « pushbroom » en contrepartie produit l'IHS en mesurant simultanément ligne par ligne dans deux dimensions : une spatiale et l'autre spectrale selon l'axe λ ; les données sont stockées sous format « BIL ». Le spectromètre ROSIS (Reflective Optics System Imaging Spectrometer) utilise le mode d'acquisition « pushbroom » [Van der Meer, 2011].

Ces deux techniques représentent des méthodes de balayage spatial.

- La technique « tunable filters » utilise des filtres variables pour acquérir une séquence d'images 2D dans différentes longueurs d'onde. L'ensemble de ces images construit le cube de données hyperspectrales ; les données sont stockées sous format « BSQ » [Liu, 2015].

Le choix du mode d'acquisition dépend de l'application considérée. En télédétection, la méthode whiskbroom est la plus utilisée [Arablouei, 2016].

Pour nos travaux de recherche, nous avons retenu les données issues des capteurs AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) et ROSIS (Reflective Optics System Imaging Spectrometer).

- AVIRIS est un spectromètre d'imagerie construit par le « Jet Propulsion Laboratory » JPL de la « National Aeronautics and Space Administration » en français l'Administration nationale de l'aéronautique et de l'espace, connue sous son acronyme (NASA) en Californie aux Etats-Unis. C'est le premier instrument de télédétection utilisé pour la caractérisation de la surface de la terre et de l'atmosphère par des mesures spectroradiométriques. Il fournit des images calibrées de la luminance spectrale dans 224 canaux (ou bandes) contiguës qui couvrent la région entre 380nm et 2500nm. Il en résulte un spectre quasi complet dans les domaines VIS-NIR-SWIR, fournissant ainsi la composition de la zone visualisée par l'instrument. La documentation technique complète sur ce capteur est disponible sur : <https://aviris.jpl.nasa.gov/aviris/index.html>.
- Le spectromètre ROSIS est un capteur aéroporté développé conjointement par « Dornier Satellite Systems » (DSS, ex-MBB), le centre aérospatial allemand (DLR) et le centre de recherche (GKSS) pour le vol sur la plateforme EURECA de l'ESA [Van der Meer, 2011]. Il est conçu pour la détection des structures spectrales fines. Dans le cadre du projet DLR HySens, la commission européenne a identifié ROSIS et un autre

capteur nommé DAIS 7915 comme étant une infrastructure de recherche majeure [Mueller, 2002]. Cet instrument utilise le mode d'acquisition « pushbroom » balayage de ligne fournissant ainsi 115 bandes spectrales contiguës. Il couvre la région spectrale entre 430nm et 860nm avec une résolution spectrale de 4nm [Riaza, 2003].

Les principales caractéristiques de ces deux capteurs sont résumées dans le tableau suivant :

Tableau 1. 1: *Caractéristiques techniques des capteurs hyperspectraux AVIRIS et ROSIS*

		AVIRIS	RODIS
Développeurs		JPL, NASA, USA	MBB/DLR/GKSS
Mode d'acquisition		Whiskbroom	Pushbroom
Domaine spectral		380nm - 2500nm	430nm – 860nm
Résolution Spatiale		20m	2m
Résolution spectrale	Nombre de bandes	224	115
	Largeur de bande	10nm	4nm

1.3.2 Prétraitement des IHS

L'analyse et l'exploitation des images hyperspectrales par les utilisateurs sont précédées par deux principaux prétraitements chez le fournisseur :

1. Les corrections radiométriques :

- L'étalonnage permet d'ajuster l'image par conversion des valeurs mesurées en réflectances absolues ou relatives [Aspinall, 2002].
- La correction atmosphérique élimine les effets perturbateurs de l'atmosphère [Brunn, 2003] tel que les nuages, la brume, la vapeur de l'eau et les aérosols qui absorbent ou diffusent les rayonnements, Figure 1.5. Ces effets dépendent du temps d'acquisition et de la zone observée. Pour les données du capteur AVIRIS de la NASA, la correction atmosphérique basée sur le transfert radiatif sans aucune mesure au sol a été proposée [Gao, 1990].

2. Les corrections géométriques :

Leur but est de corriger les distorsions dues à la position de la plateforme, sa trajectoire, son altitude ainsi que sa vitesse. Ceci nécessite l'utilisation des procédures géo-référentielles, efficaces et précises comme le PARGE [Brunn, 2003].

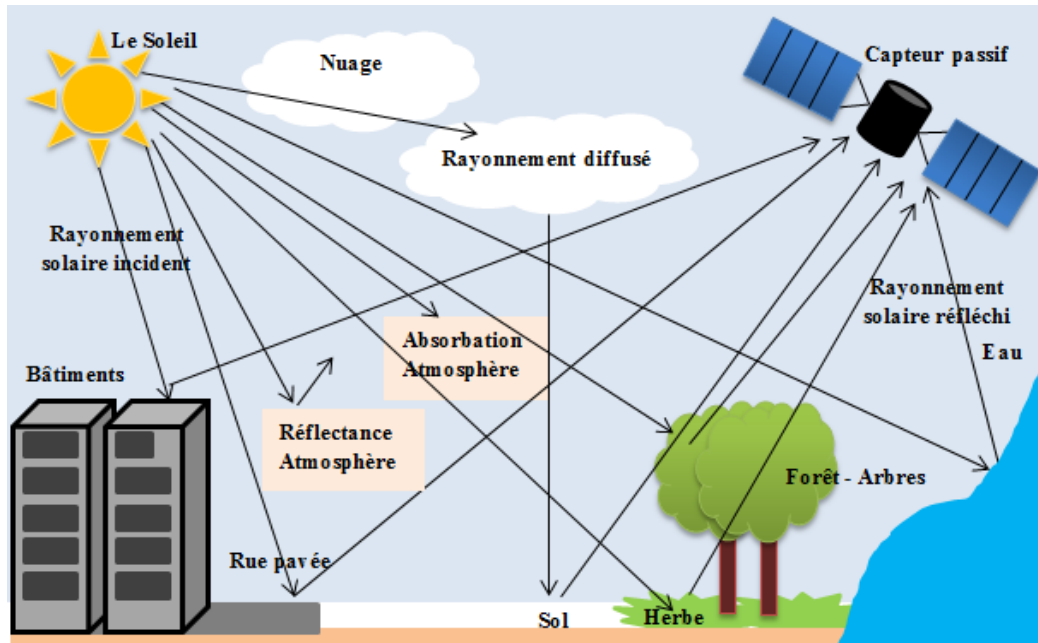


Figure 1. 5: L'imagerie hyperspectrale et les effets atmosphériques.

1.4 Problématiques relatives aux images hyperspectrales

La grande quantité d'information fournie par les IHS augmente la discrimination entre les classes et la caractérisation des couvertures terrestres. Cependant, plusieurs défis apparaissent au niveau de l'utilisation des données hyperspectrales. Ces problèmes peuvent être résumés comme suit :

- Le coût élevé du stockage et le temps de traitement.
- La variabilité spatiale des données spectrales.
- La présence du bruit dans des bandes spectrales spécifiques : il y a des bandes sensibles à la présence du brouillard, etc.
- La grande résolution spectrale avec le nombre fini des données d'entraînement qui mènent au problème de la malédiction de dimensionnalité aussi appelé phénomène de Hughes [Hughes, 1968 ; Shahshahani, 1994].

Il s'agit d'un problème commun dans la reconnaissance des formes : il y a dégradation des performances du classifieur lorsque la dimensionnalité des données d'entrée est trop élevée Figure. 1.6. Cette figure montre le taux moyen de la classification en fonction de la complexité des données d'entrée pour un nombre fini de pixels d'entraînement. D'après cette

figure, on constate que pour chaque taille des données d'entraînement, on aura un nombre maximal de bandes au-delà duquel les performances de classification se dégradent.

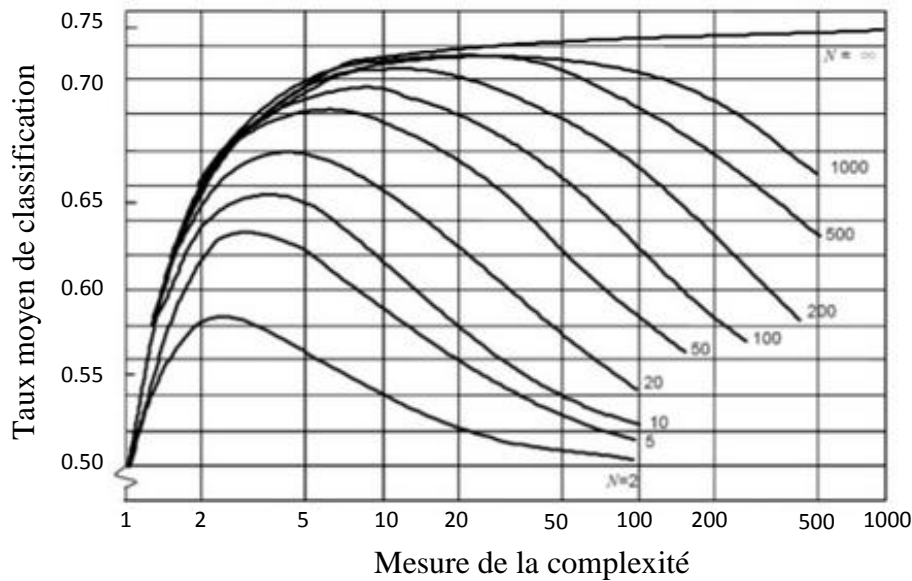


Figure 1. 6: Illustration du phénomène de Hughes [Hughes, 1968].

- La forte corrélation entre les bandes adjacentes peut également élargir les effets du phénomène précité d'où l'exigence du traitement de la redondance des données.
- Le problème du choix de la méthode de classification adéquate aussi s'impose afin de produire des cartes précises des scènes observées.

Expression de la problématique de réduction de dimensionnalité :

Lorsque la performance de classification de N bandes de l'image hyperspectrale est inférieure à la performance d'un sous-ensemble de bandes K, la réduction de dimensionnalité devient obligatoire pour la classification des images hyperspectrales [Donoho, 2000; Hasanlou, 2012; Dong, 2017]. Elle permet de réduire la quantité des informations tout en exploitant l'avantage de richesse en termes de données contenues dans les IHS. Ceci est réalisé par la sélection ou l'extraction des bandes utiles et pertinentes selon les besoins des applications [Huang, 2015 ; Medjahed, 2016].

1.5 Les images hyperspectrales d'application

Trois images hyperspectrales de différentes caractéristiques sont utilisées dans ce mémoire pour l'application des algorithmes proposés pour la réduction de dimensionnalité et la classification des IHS. Elles sont fournies par les deux capteurs AVIRIS et ROSIS.

1.5.1 AVIRIS_Indian Pines

L'image hyperspectrale AVIRIS_Indian Pines « AVIRIS92AV3C » est acquise par le capteur AVIRIS de la NASA au-dessus de la région du nord-ouest de l'Indiana en USA. Elle contient 145x145 pixels dont seulement 10366 pixels sont étiquetés de 1 à 16. Elle est composée de 224 bandes dans la région de longueur d'onde de 0,4-2,5 μm . Quatre bandes sont non informatives et donc sont enlevées. La résolution spatiale des bandes est de 20 m par pixels. Deux tiers de la scène sont couverts par des terres agricoles. Le reste contient des forêts et bâtiments. Le contenu de l'image hyperspectrale AVIRIS_Indian Pines est représenté dans la figure 1.7. Elle comprend seize classes, Tableau 1.2. AVIRIS_Indian Pines est disponible dans (<ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/92AV3C>).

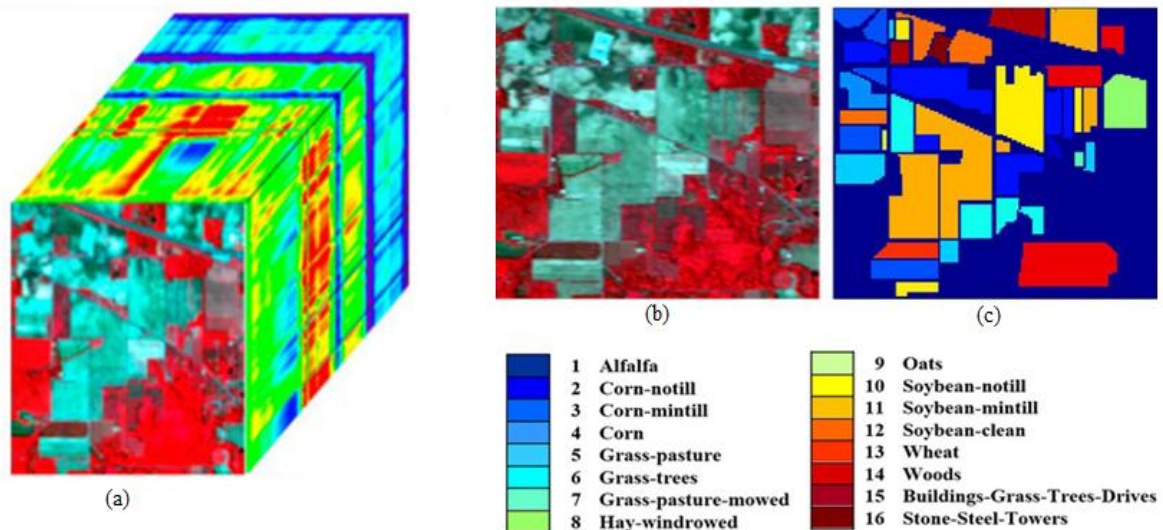


Figure 1. 7: *Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale AVIRIS_Indian Pines.*

Tableau 1. 2: Nom et Nombre de pixels de chaque classe pour l'image AVIRIS_Indian Pines

Numéro de classe	Nom de la classe en Anglais	Nom de la classe en Français	Nombre de pixels
1	Alfalfa	Luzerne	54
2	Corn-notill	Maïs non labouré	1434
3	Corn-mintill	Maïs-labourées fines	834
4	Corn	Maïs	234
5	Grass/pasture	herbe/gazon	497
6	Grass/tree	herbe/arbre	747
7	Grass/pasture-mowed	herbe/gazon tondu	26
8	Hay-windrowed	Foins- andainé	489
9	Oats	Avoine	20
10	Soybeans-notill	Soja non labouré	968
11	Soybeans-mintill	Soja-labourée fine	2468
12	Soybeans-clean	Soja-Claire	614
13	Wheat	Blé	212
14	Woods	Bois	1294
15	Building-grass-tree-drives	Construction-herbe-arbre- engins	380
16	Stone-steel towers	Tours en pierre et acier	95

1.5.2 AVIRIS_Salinas

La deuxième image utilisée est Salinas. Elle est prise par le capteur AVIRIS à la vallée de Salinas en Californie, USA. Elle se compose de 217x512 pixels dont seulement 54129 pixels sont étiquetés de 1 à 16. Elle est composée de 224 bandes de réflectance spectrale dans la région de longueurs d'onde de 0,4 à 2,5 μm . La scène de Salinas est caractérisée par une haute résolution spatiale (3,7 m pixels). Le cube hyperspectral, l'image couleur composite et la vérité de terrain correspondante de AVIRIS_Salinas sont respectivement présentés en (a), (b) et (c) de la Figure 1.8. Elle contient seize classes, Tableau 1.3. Cette IHS est connue par la mixture de ses pixels qui compliquent la séparation des classes. L'IHS AVIRIS_Salinas est disponible dans (<ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/92AV3C>).

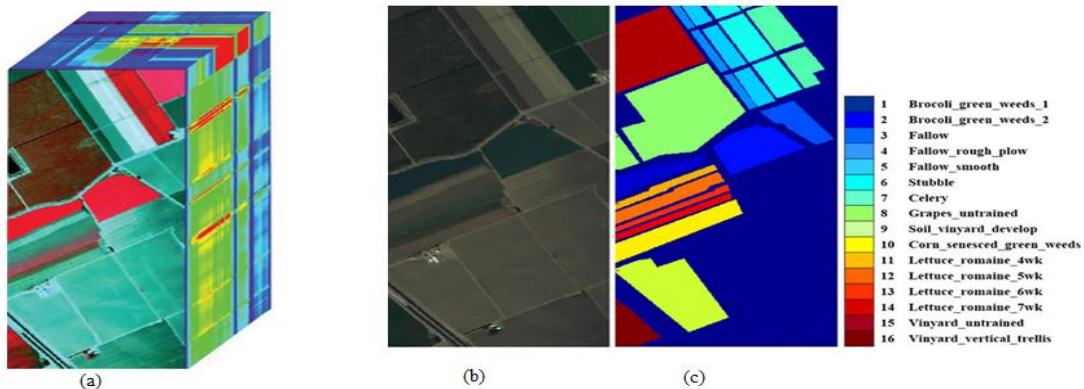


Figure 1. 8: Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale AVIRIS_Salinas.

Tableau 1. 3: *Nom et Nombre de pixels de chaque classe pour l'image AVIRIS_Salinas*

Numéro de classe	Nom de la classe en Anglais	Nom de la classe en Français	Nombre de pixels
1	broccoli green weeds1	Brocoli-mauvaises herbes vertes 1	2009
2	broccoli green weeds2	Brocoli-mauvaises herbes vertes 2	3726
3	Fallow	Jachère	1976
4	fallow rough plow	Charrue_jachère brute	1394
5	fallow smooth	Jachère lisse	2678
6	Stubble	Chaume	3959
7	Celery	céleri	3579
8	grapes untrained	raisins non taillé	11271
9	soil vineyard develop	Sol de vigne développé	6203
10	corn senesced green weeds	Maïs- mauvaises herbes vertes	3278
11	lettuce romaine 4wk	laitue romaine 4wk	1068
12	lettuce romaine 5wk	laitue romaine 5wk	1927
13	lettuce romaine 6wk	laitue romaine 6wk	916
14	lettuce romaine 7wk	laitue romaine 7wk	1070
15	vineyard untrained	Vigne non taillé	7268
16	vineyard vertical trellis	Treilles verticales de vigne	1807

1.5.3 ROSIS_Pavia University

L'image hyperspectrale ROSIS_Pavia University est une scène de 610x340 pixels recueillie par le capteur ROSIS-03 (Reflective Optics System Imaging Spectrometer) sur la zone urbaine de l'école d'ingénieurs de l'Université de Pavie en Italie. Cette zone universitaire est de faible densité contenant une grande variété de formes. Sa résolution spatiale est de 1,3 m par pixel. L'ensemble de données original comprend 115 bandes spectrales dans la plage de 0,43 à 0,86 μm où 12 bandes ont été supprimées en raison du bruit. Le cube hyperspectral, l'image en couleur composite et la vérité de terrain correspondante sont illustrés dans la Figure 1.9. Parmi les 207400 pixels de la vérité de terrain, seulement 42776 pixels (20%) sont étiquetés de 1 à 9, Tableau 1.4. L'IHS ROSIS_Pavia University est fournie par le Laboratoire de Télécommunications et Télédétection (Telecommunications & Remote Sensing Laboratory- Pr. Paolo Gamba: <http://tlclab.unipv.it/>) à l'Université de Pavie en Italie.

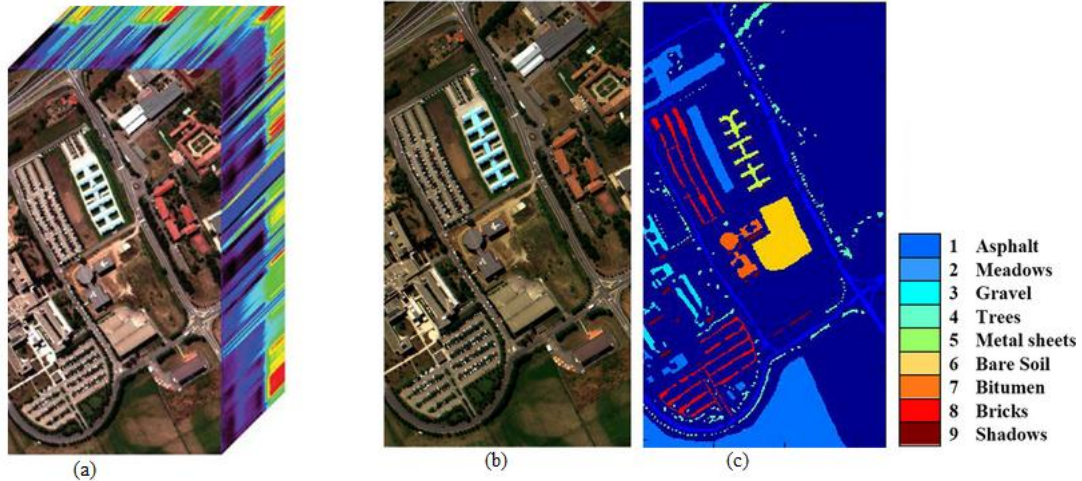


Figure 1. 9: *Le cube hyperspectral (a), l'image couleur composite (b) et la vérité de terrain correspondante (c) de l'image hyperspectrale ROSIS_Pavia University.*

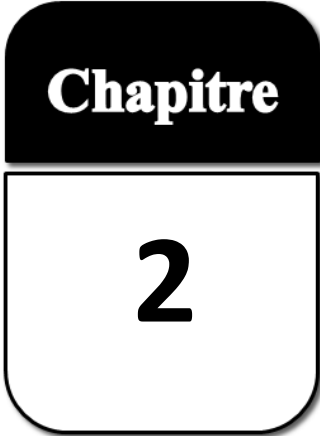
Tableau 1. 4: *Nom et nombre de pixels de chaque classe pour l'image ROSIS_Pavia University*

Numéro de classe	Nom de la classe en anglais	Nom de la classe en Français	Nombre de pixels
1	Asphalt	Asphalte	6631
2	Meadows	Prairies	18649
3	Gravel	Gravier	2099
4	Trees	Arbres	3064
5	Painted metal sheets	métal peint	1345
6	bare soil	sol nu	5029
7	bitumen	Bitume	1330
8	self-blocking bricks	briques autobloquantes	3682
9	shadows	Ombres	947

1.6 Conclusion

À travers ce premier chapitre, nous avons pu avoir un aperçu global de l'imagerie hyperspectrale, ses principales caractéristiques et applications ainsi que ses modes d'acquisition. Ceci nous a permis de dégager les défis majeurs relatifs au traitement de ces données de haute résolution.

Dans la suite de cette thèse, nous nous intéresserons au traitement de la problématique principale : la réduction de dimensionnalité et la classification des IHS. Nous proposerons de nouvelles méthodes qui permettent de réduire la dimensionnalité des images hyperspectrales par sélection et extraction d'attributs sans perdre les informations utiles. L'application sera faite sur les trois IHS retenues : AVIRIS_Indian Pines, AVIRIS_Salinas et ROSIS_Pavia University.



Chapitre
2

État de l'art des méthodes de réduction de dimensionnalité et la classification des images hyperspectrales

Résumé :

Plusieurs approches ont été élaborées pour la classification des images hyperspectrales. Leur objectif global est de classier automatiquement les pixels de l'image en classes de couverture du sol et produire des cartes thématiques. Cependant, l'énorme quantité de données disponibles nécessite un prétraitement primordial pour réduire leur dimensionnalité et par conséquent leur complexité. Ce chapitre présente une synthèse des diverses approches citées dans la littérature de classification et de réduction de dimensionnalité des images hyperspectrales.

2.1 Introduction

La majorité des applications des images hyperspectrales à la télédétection nécessitent des techniques de traitement qui permettent aux analystes du domaine d'atteindre deux objectifs fondamentaux:

- Détecter et classifier les différents objets constitutifs de chaque pixel des scènes observées.
- Réduire la dimensionnalité des données pour un traitement efficace sans perte d'informations utiles.

Dans ce chapitre, nous présenterons l'état de l'art des différentes approches qui peuvent être utilisées dans le contexte de la classification et la réduction de dimensionnalité. Les avantages et les limitations de chacune d'elles nous permettront de choisir celles qui seront adoptées dans ce mémoire pour le développement de nos propres algorithmes.

2.2 Les approches de classification des images hyperspectrales

La classification des IHS est l'un des sujets de recherche les plus intéressants en télédétection pour produire les cartes thématiques des surfaces de la terre [Lu, 2007]. Elle consiste à regrouper automatiquement tous les pixels de l'image ayant des propriétés communes en différentes classes de couverture du sol : type de sol, eau, bâtiments, végétation, route, etc.

Dans ce contexte, plusieurs algorithmes ont été développés dans la littérature et peuvent être regroupés comme étant supervisés et non supervisés, ou paramétriques et non paramétriques, ou par pixel, sous-pixel et par champ [Lu, 2007] comme indiqué dans la figure 2.1. Ces approches s'appliquent en fonction du type de données et des besoins de l'application. Une brève description de ces catégories est fournie dans ce chapitre.

2.2.1 Classification selon l'utilisation des échantillons d'entraînement

En se basant sur l'utilisation des échantillons d'apprentissage, les méthodes de classification peuvent être divisées principalement en deux catégories, à savoir les méthodes supervisées et non supervisées [Tyagi, 2008]. Les techniques de classification supervisées exigent la disponibilité d'un sous-ensemble d'échantillons de la vérité de terrain, ils sont utilisés pour l'entraînement. Les techniques non supervisées ne nécessitent aucune définition préalable des classes.

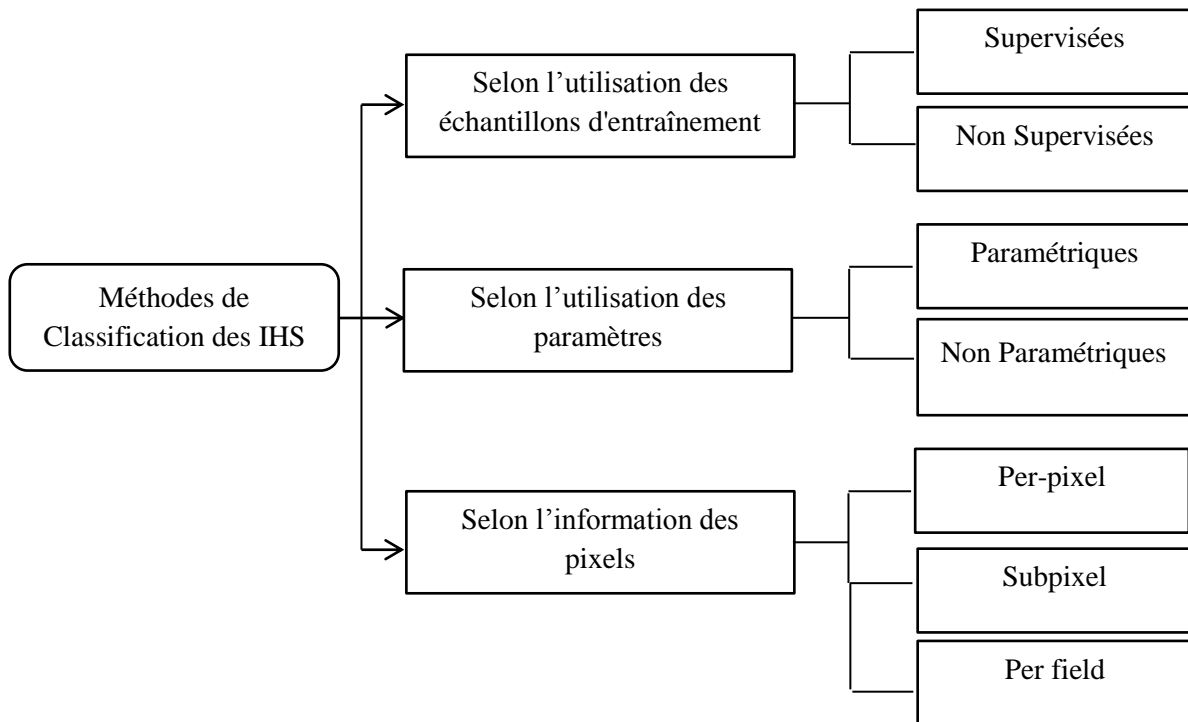


Figure 2. 1: Les approches de classification des IHS

2.2.1.1 Méthodes supervisées

La classification supervisée a pour objectif de bien estimer la fonction de mappage $Y = f(X)$ capable de prédire les variables de sortie (Y) pour les données d'entrée (X).

Dans la classification supervisée des IHS, des données de référence sont disponibles et identifiées à priori au moyen d'une combinaison de travaux sur le terrain et d'analyses cartographiques; ces caractéristiques sont utilisées comme échantillons d'entraînement pour générer des signatures qui seront utilisées ensuite pour entraîner le classifieur et classifier les données spectrales dans une carte thématique [Ablin, 2013; Lu, 2007].

Elles sont appelées supervisées parce que leur processus utilise des échantillons identifiés à priori pour attribuer les classes adéquates aux pixels inconnus. Ces prédictions sont supervisées pour être corrigées en cas d'erreur pendant l'entraînement. Ce processus d'entraînement s'arrête lorsque l'algorithme atteint un niveau de performance acceptable. La classification utilise la structure acquise pour tester les pixels inconnus.

Un grand nombre de méthodes supervisées ont été développées, les plus fréquemment utilisées sont : l'estimation du maximum de vraisemblance « Maximum likelihood estimation » [Duda, 2012], l'arbre de décision « Decision tree classifier » [Loh, 2011], le réseau de neurones « Artificial neural network » [Schalkoff, 1997], les machines à vecteurs

de support SVM [Vapnik, 1995; Kharroubi, 2002] et les forêts aléatoires « Random Forest » (RF) [Breiman, 2001; Belgiu, 2016].

2.2.1.2 Méthodes non supervisées

Pour la classification non supervisée des IHS et contrairement à l'apprentissage supervisé, aucune définition préalable des classes n'est utilisée, les données d'entrée (X) n'ont pas des étiquettes correspondantes. Dans ce cas, l'algorithme d'apprentissage regroupe automatiquement les pixels ayant des caractéristiques spectrales similaires sous forme de clusters. Ces méthodes utilisent souvent certains critères déterminés statistiquement tels que la moyenne, l'écart-type, etc.). L'analyste ré-étiquète ensuite les clusters et les combine en plusieurs classes informatives [Ablin, 2013].

Les méthodes de classification non supervisées les plus connues sont : ISODATA [Melesse, 2002] et K-means [Wagstaff, 2001].

L'algorithme ISODATA utilise une règle de décision basée sur le maximum de vraisemblance pour déterminer la classe. Il peut être réparti uniformément dans l'espace de données, puis regroupe itérativement les pixels restants à l'aide des heuristiques basées sur des distances minimales. Les pixels sont reclassés à chaque itération selon la nouvelle distance recalculée. Cette opération se poursuit jusqu'à ce qu'un nombre maximal d'itérations prédéfini ou un taux maximal de pixels ne changeant plus de catégorie.

L'algorithme K-Means est une méthode statistique utilisée pour catégoriser automatiquement un ensemble de données en k groupes. Cette méthode commence par sélectionner k centres de cluster initiaux à raffiner de manière itérative. En effet, chaque pixel est assigné au centre de cluster le plus proche qui doit être mis à jour pour être la moyenne. Cet algorithme converge, lorsqu'il n'y a plus de changement dans l'assignation des pixels aux clusters.

Les méthodes non supervisées ont donné de bons résultats dans la classification des images hyperspectrales [Niazmardi, 2013 ; Lin, 2017]. Comme elles traitent l'ensemble de l'image, elles ne sont pas sensibles au nombre d'échantillons d'entraînement, mais les relations entre les clusters et les échantillons étiquetés ne sont pas assurées. Deux problèmes majeurs sont encore rencontrés dans la classification non supervisée des images hyperspectrales: (1) comment déterminer le nombre de classes spectrales dans l'image et (2) comment trouver des échantillons qui représentent bien chacune des classes spectrales sans connaissance préalable.

Dans la littérature, il y a des méthodes dites semi-supervisées qui se situent entre l'apprentissage supervisé et non supervisé. Elles sont généralement utilisées dans le cas où vous avez une grande quantité de données d'entrée (X) avec seulement une partie de ces données est étiquetée (Y). Dans ce type de problèmes, on peut utiliser l'une des solutions suivantes :

- Utiliser des techniques de classification non supervisées pour découvrir et apprendre la structure des variables d'entrée.
- Vous pouvez également utiliser des techniques d'apprentissage supervisé pour faire des prédictions au mieux sur les données non étiquetées, puis intégrer ces données dans l'algorithme supervisé comme données d'apprentissage et utiliser le modèle pour faire des prédictions sur les nouvelles données non étiquetées.
- Un mélange de techniques supervisées et non supervisées peut aussi être adopté.

Exemple de méthodes semi-supervisées appliquées à la classification des IHS :

- C-moyennes floues « Fuzzy C-Means » [Hung, 2011; Fan, 2009].
- La carte auto-organisatrice « Self Organisation Map » [Galkowski, 2015 ; Jain, 2018].

2.2.2 Classification basée sur des paramètres

2.2.2.1 Méthodes paramétriques

Les différents algorithmes de classification font des suppositions sur la forme de la fonction de mappage. Les algorithmes qui utilisent les échantillons de données pour déterminer les paramètres statistiques tels que la moyenne, la variance, l'écart type, etc. pour classifier les pixels sont appelés méthodes paramétriques [Lu, 2007]. Ces méthodes supposent la distribution gaussienne des données.

Les algorithmes de classification paramétriques ont plusieurs avantages comme la simplicité, la rapidité et l'utilisation de moins de données d'entraînement. Cependant, ces méthodes ont des limitations. En effet, l'hypothèse d'une distribution normale est souvent violée, en particulier dans les paysages complexes où ces classifieurs produisent souvent des résultats imprécis. En outre, des échantillons d'entraînement insuffisants, ou représentatifs peuvent introduire une incertitude supplémentaire dans la classification des IHS [Lu, 2007]. Un autre inconvénient des méthodes paramétriques réside dans la complexité limitée qui ne correspond pas aux problèmes complexes rencontrés dans la pratique de l'imagerie hyperspectrale.

Des exemples des méthodes de classification paramétriques incluent: Analyse discriminante linéaire LDA [Fisher, 1936], la classification naïve bayésienne NB [Leung, 2007] et le maximum de vraisemblance « Maximum likelihood » [Duda, 2012].

2.2.2.2 Méthodes non paramétriques

Les algorithmes de classification non paramétriques sont des méthodes qui ne font pas de suppositions fermes sur la forme de la fonction de mappage. Ils n'utilisent pas de paramètres statistiques pour calculer la séparation des classes et ils peuvent s'adapter à plusieurs formes fonctionnelles. Ces algorithmes sont préférables lorsque vous avez beaucoup de données et aucune connaissance préalable [Lu, 2007, Ablin, 2013].

Parmi les approches de classification non paramétrique les plus couramment utilisées, mentionnons : les réseaux neuronaux [Schalkoff, 1997], les arbres de décision [Loh, 2011], les machines à vecteurs de support [Vapnik, 1995, Signorile, 2006] et les systèmes experts [Gallant, 1988 ; Hung, 2002].

Un autre exemple populaire de modèle non paramétrique est l'algorithme des k-voisins les plus proches qui fait des prédictions basées sur les k échantillons d'apprentissage les plus proches pour une nouvelle instance de données [Shakhnarovich, 2005].

De nombreuses recherches antérieures ont indiqué que les classifieurs non paramétriques peuvent donner de meilleurs résultats que les classifieurs paramétriques dans des paysages complexes [Delalieux, 2007 ; Yang, 2010]. Cela est dû à leur flexibilité de s'adapter à un grand nombre de formes fonctionnelles et à la performance des modèles de prédiction construits. Cependant, ces méthodes nécessitent plus de données d'entraînement pour estimer la fonction de mappage.

2.2.3 Classification basée sur l'information des pixels

Les méthodes de classification peuvent être classées, en se basant sur l'information des pixels, en classification Par pixel « Per-pixel », Sous-pixel « Subpixel » ou Par champ « Per-field ».

2.2.3.1 Méthodes par pixel 'Per-pixel'

La plupart des approches de classification sont basées sur des informations par pixel : chaque pixel est classifié dans une seule catégorie. Ces méthodes développent la signature des données en combinant les spectres de tous les pixels d'apprentissage. La signature résultante

ignore l'impact des pixels mixtes et contient la contribution de chaque matériau présent dans les pixels d'entraînement [Lu, 2007].

En raison de l'hétérogénéité des paysages des images de télédétection, les pixels mixtes sont courants dans les données à résolution spatiale moyenne et grossière. La présence de pixels mixtes a été reconnue comme un problème majeur qui affecte l'utilisation efficace des données de télédétection dans les classifications par pixel [Fisher, 1997 ; Cracknell, 1998].

La plupart des classifieurs sont basés sur l'approche per-pixel [Lu, 2007] comme: le maximum de vraisemblance [Duda, 2012], l'arbre de décision [Loh, 2011] et les machines à vecteurs de support [Vapnik, 1995].

2.2.3.2 Méthodes Sous-pixel 'Subpixel'

Les approches de classification sous-pixels consistent à fournir une estimation plus précise de la couverture terrestre. La valeur spectrale de chaque pixel est considérée comme une combinaison linéaire ou non linéaire de matériaux purs. Elles fournissent l'appartenance appropriée de chaque pixel à chacun des membres finaux. Ces méthodes sont appliquées en particulier sur les données de résolution spatiale grossières pour résoudre le problème des pixels mixtes [Foody, 1998; Woodcock, 2000]. Leur inconvénient majeur réside dans la difficulté d'évaluer la précision.

Parmi les classifieurs sous-pixels « subpixel » les plus utilisés on trouve : Les classifieurs flous « Fuzzy classifier » [Kuncheva, 2000], réseau de neurones flou [Zhang, 2001] et le modèle de mélange linéaire « Linear mixture model LMM » [Eches, 2010].

2.2.3.3 Méthodes Par-champ 'Per-field'

Dans les paysages complexes, l'hétérogénéité entraîne de fortes variations spectrales au sein d'une même classe de la couverture terrestre. Les approches de classification par pixel regroupent individuellement chaque pixel dans une certaine catégorie et les résultats peuvent être bruyants en raison de la fréquence spatiale élevée dans le paysage. Les classifieurs par champ sont conçus pour traiter ce problème de l'hétérogénéité de l'environnement et ont été efficacement appliqués pour améliorer la qualité de la classification [Aplin, 2001; Dean, 2003; Löw, 2013]. Le classifieur par champ utilise des parcelles de terrain " champs " comme unités individuelles pour calculer la moyenne du bruit [Dean, 2003]. Ces algorithmes sont essentiellement mis en œuvre par les systèmes d'informations géographiques (SIG) par l'intégration de données vectorielles et matricielles. Les données vectorielles sont utilisées

pour subdiviser l'image en parcelles qui seront ensuite utilisées pour la classification évitant ainsi les variations spectrales inhérentes à la même classe.

Cependant, la classification par champ est souvent influencée par plusieurs facteurs tels que la taille et la forme des champs, la définition des limites des champs, les propriétés spectrales, spatiales des données de télédétection et les classes de couverture terrestre choisies [Janssen, 1995]. La difficulté de traiter la dichotomie entre les modèles de données vectorielles et matricielles influe sur l'utilisation étendue de l'approche de classification par champ.

2.3 Les approches de réduction de dimensionnalité

2.3.1 Pourquoi la réduction de dimensionnalité?

Au cours des dernières années, l'énorme quantité d'informations contenues dans les images de télédétection hyperspectrales a posé de sérieux défis aux méthodes d'analyse et de classification existantes [Tang, 2014]. En effet, la présence des bandes non pertinentes ou redondantes entraîne une diminution de performances des classifieurs. Pour résoudre ce problème, les techniques de réduction de la dimensionnalité ont été introduites. Elles constituent une étape de prétraitement essentielle de la classification : cela améliore la précision et la rapidité des systèmes de prédiction [Gheyas, 2010].

Les méthodes de réduction de dimensionnalité des IHS peuvent être catégorisées selon le type de génération d'attributs et selon la dépendance au classifieur, Figure 2.2. De point de vue du processus de génération, ces approches sont subdivisées en trois procédures "Sélection", "Extraction" ou "hybride : Extraction- sélection". Concernant la dépendance au classifieur, les méthodes de réduction sont regroupées en trois catégories "Filtre", "Wrapper" et "Hybride". Une brève description de ces méthodes catégorielles est présentée dans le reste de ce chapitre.

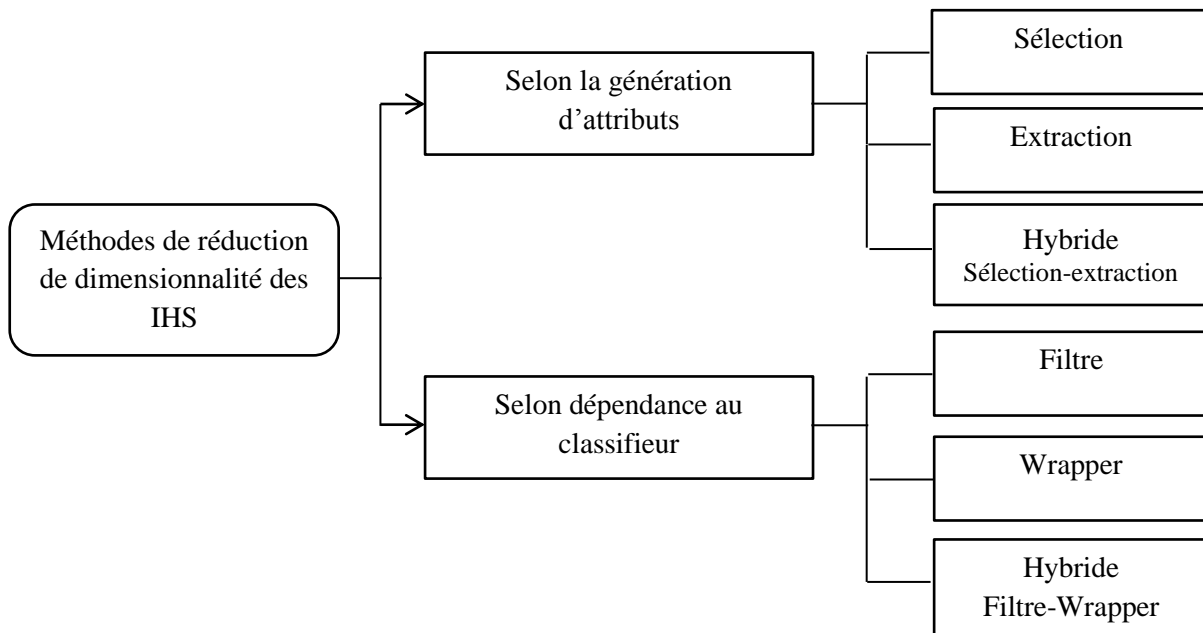


Figure 2. 2: Les approches de réduction de dimensionnalité des IHS

2.3.2 Approches de réduction selon la génération d'attributs

Les méthodes de réduction de dimensionnalité consistent à sélectionner ou extraire un sous-ensemble d'attributs de dimensionnalité K inférieur à la dimensionnalité N de l'ensemble original des données. Le sous-ensemble retenu doit permettre une observation adéquate de toutes les informations disponibles dans l'ensemble original. Ainsi, on élimine les éléments non pertinents, redondants et bruyants. Selon le processus de génération des sous-ensembles d'attributs, ces méthodes sont regroupées en deux grandes catégories: Méthodes d'extraction et de sélection. Les premières modifient l'ensemble de données d'origine pour définir un nouveau sous-ensemble de caractéristiques contenant une vaste quantité d'informations sur l'ensemble de données original. Pour les deuxièmes, un sous-ensemble des éléments originaux est identifié et sélectionné et qui couvre les informations les plus utiles parmi celles qui sont fortement corrélées et redondantes [Pal, 2010]

2.3.2.1 Méthodes basées sur la sélection d'attributs

Les méthodes basées sur la sélection d'attributs sont les plus utilisées pour réduire la dimensionnalité chez les praticiens. Elles visent à choisir un petit sous-ensemble des attributs pertinents parmi les éléments originaux en fonction de certains critères d'évaluation de la pertinence, ce qui mène habituellement à une meilleure précision de classification avec un coût moins élevé. On peut classer les méthodes de sélection selon qu'elles sont supervisées ou non, en fonction de la disponibilité des échantillons étiquetés. Si aucun échantillon étiqueté

n'est disponible, la sélection non supervisée des bandes est utilisée pour réduire la dimensionnalité : il faut identifier les bandes les plus distinctives et les plus informatives.

Les méthodes de sélection de bandes les plus utilisées en imagerie hyperspectrale sont :

- La sélection progressive de bandes (PBS) [Chang, 2014; Liu, 2017]
- La propagation par affinité (AP) [Yang, 2017]
- L'information mutuelle IM [Feng, 2015 ; Chang, 2018]
- Les algorithmes génétiques [Saqui, 2016; Nagasubramanian, 2018]

En général, les méthodes de sélection de bandes tentent de sélectionner le sous-ensemble de bandes en fonction de certains critères comme l'entropie, la variance ou tout autre critère qui préserve les propriétés physiques (l'information spectrale) de l'image hyperspectrale [Sun, 2014].

2.3.2.2 Méthodes basées sur l'extraction d'attributs

Les approches basées sur l'extraction consistent à projeter les caractéristiques dans un nouvel espace avec une dimensionnalité plus faible [Tang, 2014]. Cet espace est spécifié par un ensemble de vecteurs de caractéristiques où chaque échantillon de l'image peut être exprimé en termes de ses vecteurs générés. Parmi les exemples de techniques d'extraction de caractéristiques les plus utilisées, on trouve :

- La projection orthogonale du sous-espace (OSP) [Li, 2016; Yu, 2018]
- L'analyse en composantes principales (PCA) [Zabalza, 2014 ; Fernandez, 2016]
- L'analyse de corrélation canonique (CCA) [Luo, 2015 ; Ma, 2015].

Les méthodes de sélection et d'extraction peuvent être combinées dans des approches hybrides pour la réduction de la dimensionnalité des images hyperspectrales; Jia [Jia, 2010] propose la combinaison de la transformation en ondelettes discrètes DWT (sélection) et la propagation par affinité AP (extraction). Premièrement, les attributs issus de la transformation DWT sont extraits pour éliminer le bruit contenu dans les données hyperspectrales originales. Ensuite, l'AP est appliqué pour sélectionner les éléments représentatifs de ceux obtenus.

2.3.3 Approches de réduction selon la dépendance au classifieur

Les méthodes de réduction de dimensionnalité peuvent être classées selon la dépendance au classifieur en deux groupes : les modèles filtres et les modèles wrapper. Un autre modèle a été

déduit et appliqué par plusieurs chercheurs, il s'agit de l'approche hybride Filtre-Wrapper [Liu, 2005; Tang, 2014]. La présentation de ces différents modèles est présentée dans cette section.

2.3.3.1 Approches Filtres

Le modèle Filtre consiste à séparer le processus de sélection des attributs de l'algorithme de classification afin qu'ils n'interagissent pas. Il repose sur des mesures des caractéristiques générales des données d'apprentissage telles que la distance, la consistance, la dépendance, la mesure d'information et la corrélation [Tang, 2014].

L'approche filtre peut être schématisée comme indiqué dans la figure 2.3. La mesure de la pertinence des ensembles sélectionnés est définie indépendamment de l'algorithme d'apprentissage. Dans ce cas, la procédure de sélection des sous-ensembles est considérée comme une étape de prétraitement de la classification.

Parmi les plus importants algorithmes représentatifs du modèle filtre, on trouve:

- Relief [Robnik-Šikonja, 2003; Jia, 2013]
- Fisher score [Li, 2012 ; Jia, 2015]
- Méthodes basées sur le gain d'information [Peng, 2005; Koonsanit, 2012]

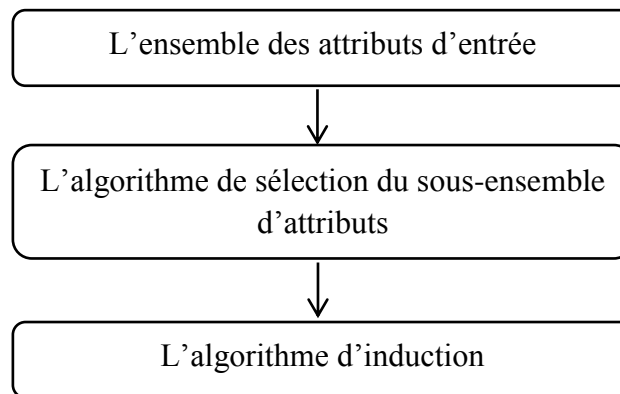


Figure 2. 3: *Principe du modèle Filtre*

2.3.3.2 Approches Wrapper

Le modèle Wrapper utilise la précision de l'algorithme d'induction prédéterminé pour juger la qualité et la pertinence des attributs sélectionnés comme indiqué dans la figure 2.4. Ces

méthodes sont extrêmement coûteuses pour traiter des données de haute dimensionnalité comportant un grand nombre d'attributs tels que les images hyperspectrales.

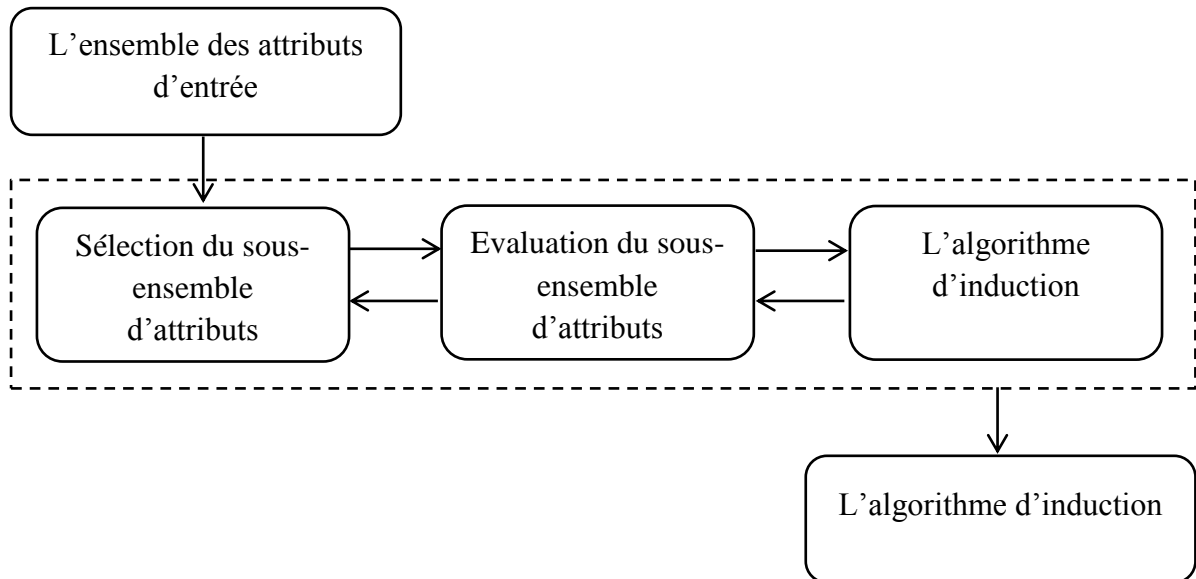


Figure 2. 4: Principe du modèle Wrapper

2.3.3.3 Approches Hybrides

L'approche Hybride représente une combinaison Filtre-Wrapper pour la sélection d'attributs. Elle a été proposée pour combler l'écart entre le modèle filtre et le modèle Wrapper [Liu, 2005]. En effet, elle intègre d'abord les critères statistiques, comme le modèle filtre, pour sélectionner plusieurs sous-ensembles candidats de caractéristiques. Ensuite, elle choisit le sous-ensemble ayant la plus grande précision de classification. Le modèle hybride effectue la sélection des attributs pendant le temps d'apprentissage. En d'autres termes, il réalise simultanément l'ajustement du modèle (model fitting) et la sélection des attributs [Cawley, 2007; Tang, 2014].

2.4 Critiques et limitations

Dans ce chapitre, nous avons distingué les algorithmes d'induction selon l'utilisation des échantillons d'apprentissage, selon le caractère paramétrique et selon le contenu statistique des pixels.

- La classification par pixel est largement utilisée dans la pratique. Cependant, la précision peut ne pas répondre aux exigences demandées en cas de présence des pixels mixtes.

- Les classifieurs non paramétriques tels que les méthodes des k plus proches voisins, l'arbre décisionnel et les machines à vecteurs de support sont les approches les plus utilisées pour la classification des données hyperspectrales.

L'étape de réduction de dimensionnalité constitue un prétraitement primordial de la classification des images hyperspectrales. Ainsi, de nombreuses méthodes ont été rapportées selon la génération des sous-ensembles réduits et l'intégration des algorithmes d'induction dans le processus de réduction. Toutes ces approches sont capables de diminuer le stockage requis, de réduire la complexité de calcul et surtout d'améliorer les performances de classification des IHS. Cette revue de littérature nous a permis de dégager les synthèses suivantes :

- L'extraction d'attributs transforme l'espace original en un nouvel espace de dimensions inférieures, souvent en combinant les caractéristiques de l'espace original. Il est difficile d'établir un lien entre les caractéristiques de l'espace original et les nouvelles caractéristiques. Par conséquent, une analyse plus poussée des nouvelles caractéristiques est problématique : souvent, il n'y a pas de signification physique pour les caractéristiques transformées obtenues.
- La sélection d'attributs vise à sélectionner un ensemble de bandes représentatives à partir de l'ensemble original et aucune nouvelle bande n'est générée. Les bandes sélectionnées par ces méthodes préservent la signification physique: elles sont exactement le sous-ensemble des bandes originales.

2.5 Conclusion

Ce chapitre présente une synthèse des diverses procédures de réduction de dimensionnalité et de la classification couramment utilisées et leur applicabilité aux données hyperspectrales. Nous avons souligné les différences entre :

- La classification supervisée, non supervisée et semi-supervisée ;
- Les méthodes paramétriques et non paramétriques ;
- La classification par pixel, sous-pixel et par champ ;

Nous avons dressé l'état de l'art des méthodes de réduction basées sur la sélection et l'extraction des attributs.

La problématique de classification et de réduction de dimensionnalité des images hyperspectrales est toujours ouverte malgré la panoplie des méthodes existantes afin d'augmenter les performances de classification et réduire le temps d'exécution.

Nous retenons, dans ce mémoire, la sélection par information mutuelle et l'extraction des caractéristiques de texture des bandes pour réduire la dimensionnalité des IHS. La classification supervisée est adoptée vu la présence des échantillons d'entraînement. Dans le chapitre suivant, les classifieurs SVM, KNN, RF et LDA seront évalués pour la classification des images hyperspectrales.

Chapitre

3

Évaluation des performances des classifieurs : SVM, RF, KNN et LDA pour la classification des images hyperspectrales en utilisant l'information mutuelle IM.

Résumé :

La classification des IHS nécessite la réduction de leur dimensionnalité comme l'indique le premier chapitre. Cette étape de classification dépend de la performance de l'algorithme d'induction (le classifieur). Dans ce chapitre, nous évaluons les performances des classifieurs supervisés suivants: les Machines à Vecteurs de Support (SVM), Random Forest (RF), les K plus proches voisins (KNN) et l'Analyse Discriminante Linéaire (LDA). Nous utilisons l'Information Mutuelle (IM) pour la réduction de dimensionnalité des IHS. Les critères retenus pour l'évaluation des performances de ces classifieurs sont : sensibilité (%), spécificité (%), Précision (%), Taux de classification globale OA (%), le coefficient Kappa et le temps d'exécution. L'application est faite sur les trois images hyperspectrales réelles, présentées dans le premier chapitre : AVIRIS_Indian-Pines, AVIRIS_Salinas et ROSIS_Pavia-University. Les résultats obtenus permettent de choisir le classifieur SVM pour la validation des algorithmes proposés, dans les chapitres suivants, pour la réduction de dimensionnalité des IHS.

3.1 Introduction

La télédétection par imagerie hyperspectrale génère des données de haute résolution spectrale. Cependant, la classification des IHS pose de nombreux problèmes tant pour l'apprentissage supervisé que pour l'apprentissage non supervisé [Janecek, 2008]. En effet, dans le cas des méthodes de classification supervisées, la malédiction de dimensionnalité ou le phénomène Hughes [Hughes, 1968 ; Bellman, 1961; Donoho, 2000] compliquent le système d'apprentissage conduisant à un mauvais modèle de classification. Cela est dû à la grande quantité d'informations spectrales avec un nombre limité d'échantillons d'apprentissage et à la présence de bandes non pertinentes et redondantes.

Dans ce chapitre, nous évaluons les performances des algorithmes de classification: les Machines à Vecteurs de Support SVM, l'Analyse Discriminante Linéaire LDA, avec différents noyaux, Random Forest RF et les K plus proches voisins KNN. La réduction de dimensionnalité des IHS est faite par sélection des bandes : nous utilisons l'information mutuelle [Sarhrouni, 2012] pour sélectionner les sous-ensembles d'images nécessaires à la classification.

3.2 Les recherches antérieures

La classification est une problématique couramment étudiée dans divers domaines et de nombreuses approches ont été développées [Lu, 2007]. Cependant, les méthodes ou algorithmes proposés dépendent des applications : ils ont des performances variées. Le choix de la méthode appropriée se fait par vérification expérimentale. Une comparaison entre les trois classifieurs SVM, KNN et Naive bayes NB a été effectuée par Bourouhou [Bourouhou, 2016]. Le but est de déterminer le classificateur le plus efficace pour la détection de la maladie de Parkinson par la voix. Les résultats obtenus concluent que le modèle basé sur les SVM est plus efficace pour la détection de cette maladie. Al Amrani [Al Amrani, 2018] compare les algorithmes de classification SVM et Random Forest RF pour l'analyse des sentiments en utilisant l'analyse de texte. Il propose aussi une technique de classification hybride entre SVM et RF nommée (RFSVM). D'après les résultats expérimentaux, il a conclu que le SVM donne des résultats mieux que RF. D'autre part, l'algorithme RFSVM semble meilleur que les deux autres : il exploite les avantages de chacune des méthodes SVM et RF.

En télédétection, Prasad [Prasad, 2017], examine la précision et la fiabilité du classifieur SVM pour la classification des images multispectrales d'Hyderabad et ses environs. Il a

comparé ses performances avec le classifieur à réseaux de neurones (ANN). Comme l'image satellitaire ne peut pas être utilisée directement par les classifieurs, ils ont introduit une technique hybride appelée « Fuzzy Incorporated Hierarchical Clustering » pour regrouper les images satellitaires multispectrales en secteurs LULC (Land Use Land Cover). Les résultats expérimentaux, sur les images fournies par le capteur multispectrale Landsat-8, montrent que le SVM produit une performance très prometteuse par rapport à l'ANN.

Camps-Valls [Camps-Valls, 2005] évalue la performance de différentes méthodes basées sur des noyaux et analyse leurs propriétés dans le contexte de la classification d'images hyperspectrales. Il considère les Réseaux Neuronaux de Fonction de Base Radiale Régulière (Reg-RBFNN), les Machines à Vecteurs de Support standard (SVMs), la méthode Kernel Fisher Discriminant KFD et AdaBoost régularisé (Reg-AB). Les expériences sont effectuées sur une portion de l'IHS AVIRIS_Indian-Pines. Il conclut que l'utilisation de SVM est plus bénéfique et donne de meilleurs résultats que les autres méthodes basées sur les noyaux, assurant une faible densité et à un coût de calcul beaucoup plus faible. L'Analyse Discriminante Linéaire LDA a été aussi introduite dans la comparaison, elle a donné la plus faible précision de classification.

Cen [Cen, 2016] effectue une étude de faisabilité d'utiliser un système d'imagerie hyperspectrale en ligne avec sélection d'attributs et classification supervisée pour détecter les dommages causés par le froid aux concombres. La comparaison est faite sur trois méthodes de sélection d'attributs spectraux (la redondance minimale de pertinence maximale MRMR, la sélection des attributs basée sur l'information mutuelle MIFS et la sélection séquentielle ascendante SFS). Trois classifieurs supervisés ont été utilisés (Naïve Bayes NB, Machines à Vecteurs de Support SVM, et K-plus Proches Voisins KNN). Les résultats indiquent que la méthode de sélection SFS en combinaison avec le classifieur SVM, donne la meilleure précision de classification que les autres combinaisons avec KNN et NB.

Dans notre étude, nous proposons de comparer les performances de quatre classifieurs supervisés (SVM, KNN, RF et LDA) en combinaison avec la sélection de bandes spectrales par information mutuelle IM dans le but d'évaluer leur application à la classification des images hyperspectrales. La principale motivation de ce choix est due aux raisons suivantes : (i) ces méthodes fournissent, dans la littérature, de bonnes performances dans de nombreuses applications dans différents domaines et notamment en télédétection [Xia, 2016; Uchino, 2012; Huang, 2016; Zhong, 2011], (ii) aucune étude n'a été faite en utilisant l'information

mutuelle pour comparer les performances de classification de ces quatre algorithmes; afin de déterminer le plus adapté pour la classification des images hyperspectrales de télédétection. Trois images hyperspectrales de différentes caractéristiques fournies par les deux capteurs hyperspectraux AVIRIS de la NASA et ROSIS sont exploitées pour l'étude comparative.

3.3 Méthodologie proposée

3.3.1 Principe

Il s'agit d'étudier et comparer les performances des quatre classifieurs (SVM, RF, LDA, et KNN) pour un choix convenable à l'application sur les images hyperspectrales. Nous adoptons la méthodologie illustrée par le schéma fonctionnel de la figure 3.1. Le processus détaillé de chaque étape de cette étude est présenté dans la suite de cette section.

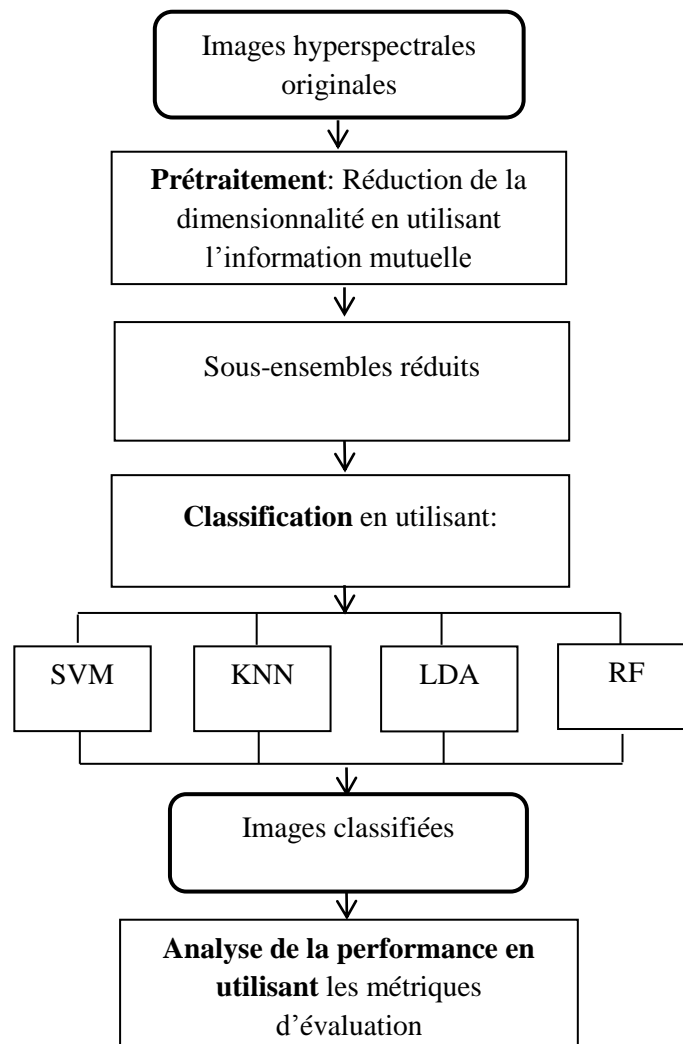


Figure 3. 1: Schéma de principe de la méthodologie.

3.3.2 Réduction de dimensionnalité des IHS par l'information mutuelle

Les images hyperspectrales fournissent plusieurs centaines de bandes de la même région. Certaines bandes sont redondantes et d'autres sont non pertinentes. La classification utilisant toutes les bandes donne des performances faibles [Hughes, 1968 ; Sarhrouni, 2012]. Il est nécessaire pour la suite de notre étude comparative des classifieurs, de réduire la dimensionnalité des IHS. Nous retenons l'information mutuelle avec une approche filtre pour générer le groupe réduit de bandes d'entrées des classifieurs [Sarhrouni, 2012].

3.3.2.1 Mesure de l'information mutuelle

Cette section présente les principes de la théorie de l'information en mettant l'accent sur l'entropie et l'information mutuelle et leur utilisation en sélection d'attributs.

L'entropie d'une variable aléatoire X est notée $H(X)$: c'est la quantité d'information nécessaire pour décrire cette variable [Cover, 2012]. Soit $p(x)$ la fonction de densité de probabilité de X et $p(y)$ la fonction de densité de probabilité de Y .

$$H(X) = -\sum_{x \in X} p(x) \log(p(x)) \quad (3.1)$$

$$H(Y) = -\sum_{y \in Y} p(y) \log(p(y)) \quad (3.2)$$

On définit l'entropie conjointe $H(X, Y)$ des deux variables X et Y :

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log(p(x, y)) \quad (3.3)$$

Où $p(x, y)$ est la fonction de densité de probabilité conjointe des variables X et Y .

L'information mutuelle (IM) est la quantité d'information que les deux variables X et Y partagent. Le diagramme de Venn suivant illustre les relations décrites entre les entropies et l'information mutuelle.

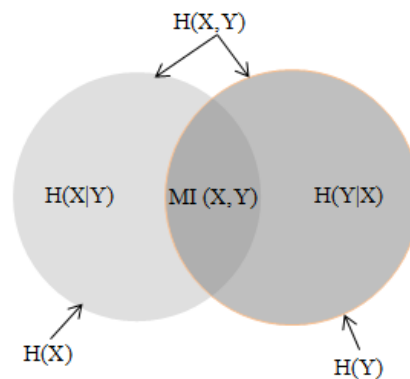


Figure 3. 2: Diagramme de Venn de la relation entre les entropies et l'IM de X et Y .

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \quad (3.4)$$

$$MI(X, Y) = MI(Y, X) \quad (3.5)$$

Comme défini à l'équation (3.4), l'indicateur IM calcule alors la différence entre les entropies indépendantes et l'entropie conjointe pour estimer la dépendance statistique des deux variables X et Y. L'IM est nulle si les variables aléatoires sont statistiquement indépendantes.

De nombreuses méthodes basées sur l'IM ont été proposées pour la sélection d'attributs dans divers problèmes de reconnaissance de formes, notamment en télédétection [Viola, 1997 ; Guyon, 2003 ; Guo, 2006 ; Sarhrouni, 2013 ; Battiti, 1994 ; Kwak, 2002 ; Huang, 2006 ; Martinez, 2007]. Elles consistent à maximiser la pertinence tout en minimisant la redondance. Dans la suite de cette étude, nous nous basons sur l'algorithme Sarhrouni [Sarhrouni, 2012] dont nous étendrons l'application sur deux autres IHS.

3.3.2.2 Présentation de la sélection des bandes spectrales basée sur l'IM

L'équation 3.4 est utilisée pour calculer l'IM entre la vérité de terrain nommée X et chaque bande de l'IHS nommée Y.

Le défi de la plupart des méthodes de sélection basées sur l'information mutuelle réside dans le traitement de la redondance entre les bandes sélectionnées. En effet, les bandes choisies pour leurs valeurs de MI élevées peuvent contenir la même information qui n'ajoute rien au résultat de la classification. Pour résoudre ce problème, l'algorithme [Sarhrouni, 2012] consiste à construire à chaque fois une vérité de terrain approximative nommée GT-est en utilisant les bandes déjà sélectionnées à cette étape par calcul de leur moyenne. Chaque bande candidate est retenue si elle augmente l'information mutuelle entre GT-est et la Vérité de terrain GT. Ces étapes sont résumées comme suit :

- 1) Calculer l'information mutuelle MI entre la vérité du terrain GT et chaque bande de l'image hyperspectrale originale.
- 2) Ordonner les bandes par valeur décroissante de leurs IM avec la vérité de terrain.
- 3) Initialiser la vérité de terrain approximative (GT-est) par la bande à plus grande valeur d'IM avec la vérité de terrain.
- 4) Pour toute bande candidate B_i , elle est conservée si elle augmente la dernière valeur l'IM de la GT et GT-est par un seuil de contrôle de redondance.

3.3.3 Classifieurs étudiés

Dans cette section, nous présentons le principe des quatre algorithmes de classification supervisés SVM, RF, KNN et LDA. Chaque classifieur réalise une étape d'apprentissage pour identifier un modèle décrivant la relation entre les réflectances spectrales des pixels d'entraînement et leurs labels de classe. La variété des catégories des classifieurs choisis est telle que : les quatre classifieurs sont par pixel et supervisés ; le SVM, RF et le KNN sont non paramétriques, mais le LDA est paramétrique, comme présenté au chapitre 2.

3.3.3.1 Machines à Vecteurs de Support SVM

Les Machines à Vecteurs de Support ou en anglais « Support Vector Machines (SVM) est un paradigme d'apprentissage à deux classes introduit par Vapnik [Vapnik, 1995], c'est un classifieur supervisé et non-paramétrique [Lu, 2007]. Il a été utilisé dans la classification des données hyperspectrales de télédétection et a donné de bons résultats dans de nombreux travaux [Mountrakis, 2011; Xie, 2017].

Nous nous sommes référés à Kharroubi [Kharroubi, 2002] et Signorile [Signorile, 2006] pour décrire le principe général du classifieur SVM : Il consiste à projeter les données de l'espace d'entrée, qui sont non linéairement séparables, vers un espace de plus grande dimension nommé espace de caractéristiques. L'objectif est de rendre les données linéairement séparables. Un hyperplan optimal qui sépare les classes est construit dans cet espace à condition que:

- Les vecteurs appartenant aux différentes classes se situent de différents côtés de l'hyperplan.
- La maximisation de la marge M qui est la plus petite distance entre les vecteurs et l'hyperplan.

✓ Pour les données linéairement séparables

Il existe une infinité d'hyperplans qui peuvent séparer parfaitement les données linéairement séparables. Soit $h(x)=w.x+b$. Un hyperplan qui satisfait les conditions suivantes :

$$\begin{cases} w \cdot x_i + b \geq 1 & \text{si } y_i = 1 \\ w \cdot x_i + b \leq -1 & \text{si } y_i = -1 \end{cases} \quad (3.6)$$

L'hyperplan optimal est celui qui maximise la marge M , c'est équivalent à maximiser la somme des distances des classes par rapport à l'hyperplan. Ce principe est exprimé comme suit :

$$M = \min_{(x_i|y_i=1)} \frac{w \cdot x + b}{\|w\|} - \max_{(x_i|y_i=-1)} \frac{w \cdot x + b}{\|w\|} = \frac{1}{\|w\|} - \frac{-1}{\|w\|} = \frac{2}{\|w\|} \quad (3.7)$$

Maximiser $\frac{2}{\|w\|}$ est équivalent à minimiser $\frac{\|w\|^2}{2}$ sous la contrainte $y_i(w \cdot x_i + b) \geq 1$ pour $i \in 1, \dots, m$

Ceci mène à l'hyperplan optimal dont l'équation est :

$$h(x) = w \cdot x + b_0 = \sum_{i=1}^m \alpha_i y_i (x \cdot x_i) + b_0 \quad (3.8)$$

Où : $\alpha_i \geq 0$, avec maximisation de $\sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$ sous la contrainte $\sum_{i=1}^m \alpha_i y_i = 0$.

✓ **Pour les données non linéairement séparables :**

Dans les applications réelles, il est rare que les classes soient linéairement séparables [Signorile, 2006]. L'hyperplan optimal des données non linéairement séparables doit satisfaire les conditions suivantes [Kharroubi, 2002] :

- Maximisation de la distance entre les vecteurs bien classés et l'hyperplan optimal
- Minimisation de la distance entre les vecteurs mal classés et l'hyperplan optimal

Ceci nécessite l'introduction des variables de pénalité non négatives ξ_i , appelées variables d'écart, pour $i \in 1, \dots, m$. La contrainte devient :

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \quad \text{pour } i \in 1, \dots, m \quad (3.9)$$

L'objectif est la minimisation de la fonction suivante :

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \quad (3.10)$$

Avec $C \geq 0$, c'est le paramètre de régularisation.

Pour s'affronter aux limitations de cette méthode de séparation linéaire, des fonctions noyau ont été introduites [Cornuéjols, 2002]. Dans l'équation de l'hyperplan optimal 3.8, le produit scalaire $(x_i \cdot x)$ peut être remplacé par une fonction noyau. L'hyperplan devient :

$$h(x) = w \cdot x + b_0 = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b_0 \quad (3.11)$$

Il existe plusieurs noyaux, dans cette étude nous utilisons :

- Noyau linéaire:

$$K_{lin}(x_i, x) = (x \cdot x_i) \quad (3.12)$$

- Noyau Fonction de Base Radiale « Radial Basis Function » (RBF):

$$K_{RBF}(x_i, x_j) = \exp(-\gamma \|x - x_i\|^2) \quad (3.13)$$

- Noyau sigmoïde

$$K_{sig}(x_i, x_j) = \tanh(x \cdot x_i + c) \quad (3.14)$$

✓ SVM multi-classes

Le SVM a été initialement conçu pour résoudre des problèmes de classification à deux classes. Pour l'imagerie hyperspectrale de télédétection, nous avons plusieurs classes, l'évaluation s'étend alors aux SVM multi-classes. Deux principales approches ont été proposées pour y parvenir [Keshava, 2001 ; Hsu, 2002]. Pour un problème de k classe, nous pouvons utiliser « un contre tous » (one versus all) ou « un contre un » (one versus one) que nous avons choisi dans notre étude. Ainsi nous exécutons $\frac{k(k-1)}{2}$ classifieurs binaires et on attribue à un élément la classe majoritaire parmi les $\frac{k(k-1)}{2}$ de fonction de décision [Hsu, 2002].

3.3.3.2 K- plus proches Voisins K-NN « K-Nearests Neighbors »

Le classifieur K-plus proches Voisins (KNN) ou en anglais « K-Nearest Neighbor » est un algorithme d'apprentissage non paramétrique ; c'est l'un des plus utiles en tant que classifieur supervisé basé sur la mesure de similarité. Il a été appliqué avec succès à la classification des images hyperspectrales [Li, 2015; Cariou, 2016].

Soit X les données d'entraînement de n dimensions. X est étiquetées parmi les C classes et X* sont les nouveaux points de test dont on doit vérifier la classe:

$$X = \{x_1, x_2, x_3, \dots, \dots, x_n\} \quad (3.15)$$

$$C = \{c_1, c_2, c_3, \dots, \dots, c_n\} \quad (3.16)$$

$$X^* = \{x^*_1, x^*_2, x^*_3, \dots, \dots, x^*_n\} \quad (3.17)$$

La procédure générale du classifieur KNN peut être résumée dans ces étapes :

- Trouver la distance entre chaque point des données d'entraînement x_i et le nouveau point X^* . Dans cette étape, différentes métriques peuvent être utilisées pour déterminer cette distance tel que : Euclidienne, euclidienne standardisée, mahalanobis, cityblock, distance chebychev [Chomboon, 2015]. La distance la plus populaire est la distance

euclidienne que nous utilisons dans cette étude donnée, dans un espace Euclidien de n dimensions, par la formule suivante:

$$d(x, x^*) = \sqrt{(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2 + \dots + (x_n - x_n^*)^2} \quad (3.18)$$

- Trier ces distances par ordre croissant.
- Retourner les k points en x_i qui sont les plus proches du nouveau point X^* .
- Obtenir la classe la plus fréquente de ces points.

Pour $k=1$: Le cas de l'algorithme du plus proche voisin (1NN), le nouveau point reçoit l'étiquette de classe du plus proche voisin.

Pour $k>1$: Le cas de K-plus proches voisins (K-NN), Le nouveau point est classé en votant le voisin le plus fréquent.

Le choix de la valeur optimale de K est critique. En général, la valeur de précision de classification augmente avec une valeur élevée de K parce qu'elle réduit le risque de fausse classification, mais le coût de calcul augmente aussi. Dans cette étude, nous implémentons le classifieur KNN en utilisant différentes valeurs de 1 à 13. Nous retenons les valeurs de K : 1, 3, 5 et 7 pour raison de stagnation des performances.

3.3.3.3 Analyse discriminante linéaire LDA

L'analyse discriminante linéaire LDA est également connue sous le nom d'analyse discriminante de Fisher, nommé en référence à son développeur R.A. Fisher [Fisher, 1936]. Lorsqu'il est appliqué à la classification des images hyperspectrales, cet algorithme vise à trouver une combinaison linéaire des caractéristiques pour séparer les classes de l'image. En effet, cette approche maximise le rapport de la variance entre les classes et celle à l'intérieur d'une classe pour assurer une séparabilité maximale [Balakrishnama, 1998]. Dans cette étude, nous avons utilisé deux types noyaux de LDA qui sont linéaires et diag-linéaires où:

- Linéaire : (par défaut) estime une matrice de covariance pour toutes les classes.
- Diag-linéaire : utilise la diagonale de la matrice de covariance linéaire.

Nous nous référons à Raschka [Raschka, 2014] pour résumer le principe du LDA dans les étapes suivantes :

- Pour chaque classe, calculez les vecteurs moyens μ_j et le vecteur moyen de l'ensemble des données nommées μ_e :

$$\mu_e = \sum p_j \times \mu_j \quad (3.19)$$

Avec p_j sont les probabilités priori des classes.

- Pour formuler le critère de séparation des classes, calculez la dispersion à l'intérieur de la classe « within-class scatter » et entre les classes « between class scatters »:

Pour la dispersion à l'intérieur de la classe, nous utilisons l'équation :

$$S_w = \sum_j p_j \times (cov_j) \quad (3.20)$$

Avec j est le nombre des classes, et : $cov_j = (x_j - \mu_j)(x_j - \mu_j)^T$

Pour la dispersion entre les classes, nous utilisons l'équation suivante:

$$S_b = \sum_j (\mu_j - \mu_e)(\mu_j - \mu_e)^T \quad (3.21)$$

- Pour les matrices de dispersion, calculer les vecteurs propres (e_1, e_2, \dots, e_n) et leurs valeurs propres correspondantes ($\lambda_1, \lambda_2, \dots, \lambda_n$).

- Sélectionner les k vecteurs propres ayant les plus grandes valeurs propres pour former la matrice de dimension $n \times k$ appelée W .

- Transformer les échantillons dans un nouveau sous-espace en utilisant la matrice W et la distance (euclidienne ou la valeur quadratique moyenne quadratique (RMS), etc.) pour classer les points de données ; la distance euclidienne est calculée selon l'équation suivante:

$$dis_n = W^T \times X - m_{ntrans} \quad (3.22)$$

Avec m_{ntrans} est la moyenne de l'ensemble de données transformé et n le nombre de classes.

3.3.3.4 Les forêts aléatoires « Random Forest RF »

Les forêts aléatoires (RF) [Breiman, 2001] est l'un des algorithmes d'apprentissage couramment utilisés dans de nombreux types de problèmes de science des données [Belgiu, 2016; Al Amrani, 2018]. Comme son nom l'indique, ce classifieur consiste à construire une multitude d'arbres de décision (DCT) pour l'entraînement. Son idée principale est qu'un groupe de classifieurs faibles puissent se réunir pour former un classifieur fort.

Les étapes générales pour réaliser l'algorithme RF sont énumérées ci-dessous :

- Construisez la forêt aléatoire :
 - Sélectionnez " m " caractéristiques parmi les caractéristiques totales " X ".
 - Pour $m < X$: Calculer les nœuds-parents " n " et leurs sous-nœuds "partitions filles" en utilisant le meilleur point de division.

- Répétez ces deux étapes jusqu'à ce que vous ayez des nœuds "d" et la cible comme nœud feuille.
 - Répétez les étapes précédentes jusqu'à ce que vous ayez une forêt avec des arbres "k".
- La prédiction de la forêt aléatoire.
- Stockez les résultats prédits en utilisant les arbres de décision créés et les meilleures caractéristiques.
 - Utilisez le vote majoritaire pour chaque cible prévue.
 - La prédiction finale de l'algorithme RF est la cible prédite votée la plus élevée.

Dans l'algorithme RF, les arbres doivent être définis en respectant un bon choix de l'arborescence appropriée, les caractéristiques des sous-ensembles et la stratégie de règles de décision à utiliser à chaque nœud interne [Safavian, 1991].

3.4 Résultats et discussion

3.4.1 Description des bases de données hyperspectrales utilisées

Pour les expériences, nous utilisons les trois images hyperspectrales réelles présentées dans le premier chapitre : AVIRIS_Indian Pines, AVIRIS_Salinas et ROSIS_Pavia University. Ces IHS ont des caractéristiques différentes en termes de nombre de bandes, de classes et de type de caractéristiques. Les figures 3.3, 3.4 et 3.5 représentent les vérités de terrain et des exemples de bandes pour respectivement AVIRIS_Indian-Pines, AVIRIS_Salinas et ROSIS_Pavia-University.

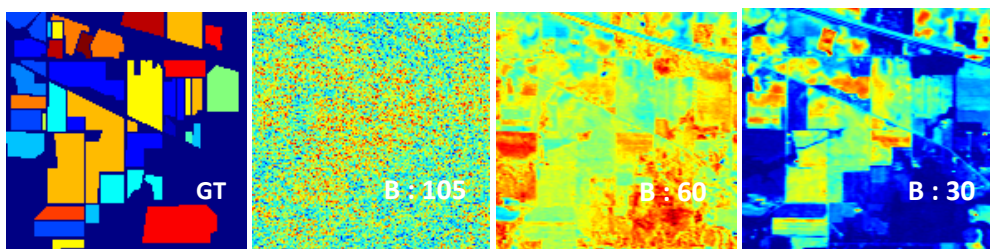


Figure 3. 3: Exemples de bandes d'AVIRIS Indian Pines avec la vérité de terrain correspondante.

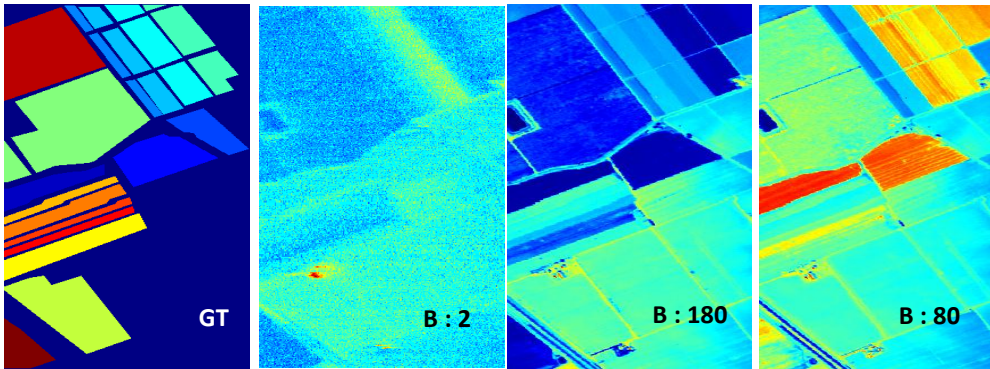


Figure 3. 4: Exemples de bandes d'AVIRIS_Salinas avec la vérité de terrain correspondante.

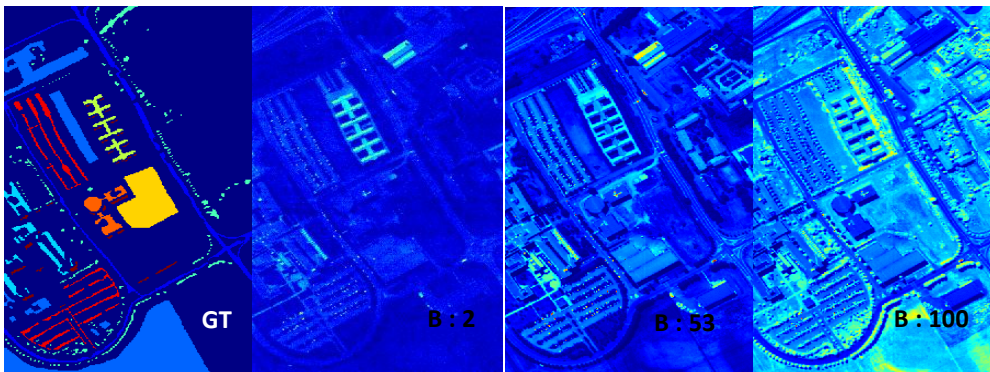


Figure 3. 5: Exemples de bandes de ROSIS_Pavia-University avec la vérité de terrain correspondante.

3.4.2 Paramètres de classification et d'évaluation

Pour classifier les pixels des images précitées à l'aide des quatre classifieurs supervisés, 50% des pixels libellés de chaque classe sont choisis au hasard comme données d'entraînement et le reste est utilisé pour le test. Les bandes sélectionnées à l'aide de l'information mutuelle sont utilisées comme données d'entrée pour les classifieurs. Les valeurs des paramètres de chaque classifieur sont choisies expérimentalement en fonction de la plus grande précision de classification. Toutes les expériences sont compilées sur un ordinateur quadri-cœur Duo, 64-b, à fréquence de CPU 2.1 Ghz avec 3 Go de RAM.

Afin de comparer les performances des classifieurs, différents coefficients d'évaluation sont calculés à partir de la matrice de confusion. Parmi les plus utilisées dans la littérature, nous avons choisi la sensibilité, la spécificité et la précision calculées selon les formules suivantes :

$$\text{Sensibilité} = \frac{VP}{VP+FN} \quad (3.23)$$

$$\text{Spécificité} = \frac{VN}{VN+FP} \quad (3.24)$$

$$Précision = \frac{VP}{VP+FP} \quad (3.25)$$

Où VP est vrai positif, VN est vrai négatif, FP est faux positif et FN est faux négatif.

Deux autres mesures populaires largement utilisées en télédétection par imagerie hyperspectrale pour la comparaison ont été utilisées dans cette étude, à savoir la précision globale (OA) qui représente les pixels correctement classés sur tous les échantillons testés. Le coefficient kappa (k) est utilisé pour mesurer le degré de concordance [Cohen, 1960]. Le temps de calcul est également utilisé. Le coefficient de concordance kappa est formulé comme suit :

$$k = \frac{N * \sum_{i=1}^n m_{i,i} - \sum_{i=1}^n A_i B_i}{N^2 - \sum_{i=1}^n A_i B_i} \quad (3.26)$$

Avec:

i est le numéro de classe, N est le nombre total de valeurs classifiées.

$m_{i,i}$ est le nombre de valeurs correctement classées dans la classe i .

A est le nombre des vraies valeurs de la classe i et B est le nombre de valeurs prédites de la classe i .

3.4.3 Résultats et analyses

La première étape dans cette étude consiste à sélectionner les bandes pertinentes à partir des ensembles de données originaux qui contiennent 224 bandes pour Indian-Pines et Salinas et 103 bandes pour Pavia-University. Cette étape a été réalisée avec la méthode filtre basée sur l'information mutuelle [Sarhrouni, 2012]. Par la suite, nous avons appliqué les quatre classifieurs supervisés avec des noyaux différents sur les sous-ensembles de bandes sélectionnées.

3.4.3.1 Comparaison des taux de classification des algorithmes selon le nombre de bandes sélectionnées

Les résultats expérimentaux obtenus sur les trois images hyperspectrales Indian-Pines, Salinas et de Pavia-University sont résumés et présentés respectivement aux tableaux 3.1, 3.2 et 3.3 et dans les figures 3.6, 3.7 et 3.8.

Tableau 3. 1: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas d'Indian-Pines.

		SVM			RF	LDA		KNN			
		SVM-RBF	SVM-Linéaire	SVM-Sigmoid		LDA-Linéaire	LDA-Diag Linéaire	KNN-k=1	KNN-k=3	KNN-k=5	KNN-k=7
Nombre de bandes sélectionnées	5	51,82	47,2	46,7	50,94	34,72	32,94	41,71	45,31	47,54	48,98
	10	55,43	48,35	46,64	54,26	38,32	32,96	44,63	47,87	50,56	51,90
	15	57,65	49,27	46,68	56,5	40,41	33,12	48,08	51,02	52,93	54,51
	20	63,08	54,51	46,62	60,02	44,63	33,92	53,56	56,22	58,21	58,82
	25	66,12	59,77	46,7	61,52	49,44	33,84	55,93	58,35	59,59	60,61
	30	73,54	67,52	46,74	68,38	54,88	34,42	63,47	65,83	67,33	66,84
	35	76,06	69,08	46,64	70,54	56,34	35,05	66,94	69,22	71,07	71,07
	40	78,96	70,82	46,74	72,98	58,99	35,69	72,55	73,76	73,49	73,47
	45	80,85	71,48	46,72	73,86	60,55	36,14	73,95	75,39	75,56	74,79
	50	81,63	71,6	46,74	73,8	61,85	36,35	74,95	76,21	75,92	75,55
	55	82,06	72,39	46,56	75,00	62,32	36,59	75,92	76,78	76,66	76,17
	60	82,74	72,67	46,35	74,71	63,04	37,11	76,04	77,01	77,50	76,31
	65	86,38	75,63	46,35	78,32	64,68	37,62	80,97	81,47	81,49	80,44
	70	86,95	76,66	46,13	79,52	65,65	38,22	82,21	82,23	81,59	81,08
	75	86,81	76,76	45,94	79,12	66,69	38,01	82,10	82,15	81,47	81,02
	80	87,28	77,69	45,9	79,76	68,21	38,73	81,92	82,31	81,43	80,95
	85	88,31	79,91	45,86	79,99	71,09	38,85	81,78	82,06	81,59	80,77
90	87,75	80,11	45,88	79,97	71,26	38,96	80,21	80,54	80,52	80,25	
95	87,06	80,07	45,9	79,6	71,63	38,91	78,30	79,49	79,80	79,58	
100	85,62	79,51	45,86	79,91	71,67	38,79	73,82	75,88	76,21	75,96	

Tableau 3. 2: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas de Salinas

		SVM			RF	LDA		KNN			
		SVM-RBF	SVM-Linéaire	SVM-Sigmoid		LDA-Linéaire	LDA-Diag Linéaire	KNN-k=1	KNN-k=3	KNN-k=5	KNN-k=7
Nombre de bandes sélectionnées	5	81,33	76,07	48,95	86,70	70,53	60,26	84,30	82,04	81,67	81,69
	10	82,69	77,75	49,96	88,32	72,04	58,26	84,95	82,66	82,26	82,08
	15	85,64	82,55	50,13	89,69	77,76	59,79	86,91	84,59	84,32	84,11
	20	86,51	83,82	50,12	90,14	78,42	58,56	87,25	85,43	85,02	84,86
	25	86,88	84,3	50,08	90,24	78,64	58,08	87,19	85,44	85,12	85,01
	30	87,56	85,28	49,78	90,71	79,86	57,97	87,42	85,66	85,50	85,41
	35	88,42	86,47	49,25	91,07	81,12	58,00	87,88	86,05	85,98	85,92
	40	88,97	87,28	48,55	91,43	82,01	58,15	88,35	86,54	86,35	86,27
	45	91,11	89,25	52,54	93,11	86,62	67,83	91,08	89,85	89,63	89,33
	50	91,14	89,77	52,02	93,29	87,87	68,92	91,55	90,10	89,74	89,43
	55	91,83	89,89	51,89	93,26	88,13	69,33	91,41	90,11	89,70	89,62
	60	91,85	89,87	51,87	93,34	88,18	69,75	91,57	90,16	89,87	89,79
	65	92,04	90,05	51,80	93,67	88,16	70,89	91,76	90,43	90,15	89,91
	70	92,86	91,08	51,73	94,03	89,07	71,66	91,86	90,63	90,31	90,23
	75	93,11	91,1	51,65	94,20	89,32	72,08	92,10	90,72	90,46	90,21
	80	93,2	91,15	51,56	94,18	89,51	72,64	92,13	90,76	90,67	90,28
	85	93,27	91,23	51,53	94,23	89,67	73,33	92,21	90,95	90,60	90,28
	90	93,26	91,23	51,53	94,23	89,64	73,33	92,21	90,95	90,60	90,28
95	93,26	91,23	51,53	94,17	89,64	73,33	92,21	90,95	90,60	90,28	
100	93,26	91,23	51,53	94,15	89,64	73,33	92,21	90,95	90,60	90,28	

Tableau 3. 3: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon le nombre des bandes sélectionnées dans le cas de Pavia-University

		SVM			RF	LDA		KNN			
		SVM-RBF	SVM-Linéaire	SVM-Sigmoid		LDA-Linéaire	LDA-Diag Linéaire	KNN-k=1	KNN-k=3	KNN-k=5	KNN-k=7
Nombre de bandes sélectionnées	5	67,15	65,32	62,07	64,98	46,5	46	57,17	60,7	63,01	64,1
	10	74,29	71,4	61,63	74,38	57,97	47,54	66,94	70,05	71,76	72,58
	15	78,68	75,8	61,04	79,13	63,22	51,65	71,75	75,21	76,50	77,60
	20	85,06	80,72	60,12	84,96	68,48	50,2	79,5	82,12	83,16	83,56
	25	89,08	84,53	58,77	88,86	72,23	48,88	85,39	87,51	87,89	88,08
	30	90,58	85,19	58,2	90,45	73,01	49,11	87,39	89,22	89,62	89,68
	35	91,65	85,53	76,5	90,88	75,4	54,63	86,41	87,95	88,10	88,28
	40	91,65	85,53	76,5	90,91	75,4	54,63	86,41	87,95	88,1	88,28
	45	91,65	85,53	76,5	91,02	75,4	54,63	86,41	87,95	88,1	88,28
	50	91,65	85,53	76,5	90,98	75,4	54,63	86,41	87,95	88,1	88,28
	55	91,65	85,53	76,5	91,01	75,4	54,63	86,41	87,95	88,1	88,28
	60	91,65	85,53	76,5	91,07	75,4	54,63	86,41	87,95	88,1	88,28
	65	91,65	85,53	76,5	91,02	75,4	54,63	86,41	87,95	88,1	88,28
	70	91,65	85,53	76,5	91,06	75,4	54,63	86,41	87,95	88,1	88,28
	75	91,65	85,53	76,5	90,95	75,4	54,63	86,41	87,95	88,1	88,28
	80	91,65	85,53	76,5	91,08	75,4	54,63	86,41	87,95	88,1	88,28
	85	91,65	85,53	76,5	91,06	75,4	54,63	86,41	87,95	88,1	88,28
90	91,65	85,53	76,5	91,06	75,4	54,63	86,41	87,95	88,1	88,28	
95	91,65	85,53	76,5	91,06	75,4	54,63	86,41	87,95	88,1	88,28	
100	91,65	85,53	76,5	91,06	75,4	54,63	86,41	87,95	88,1	88,28	

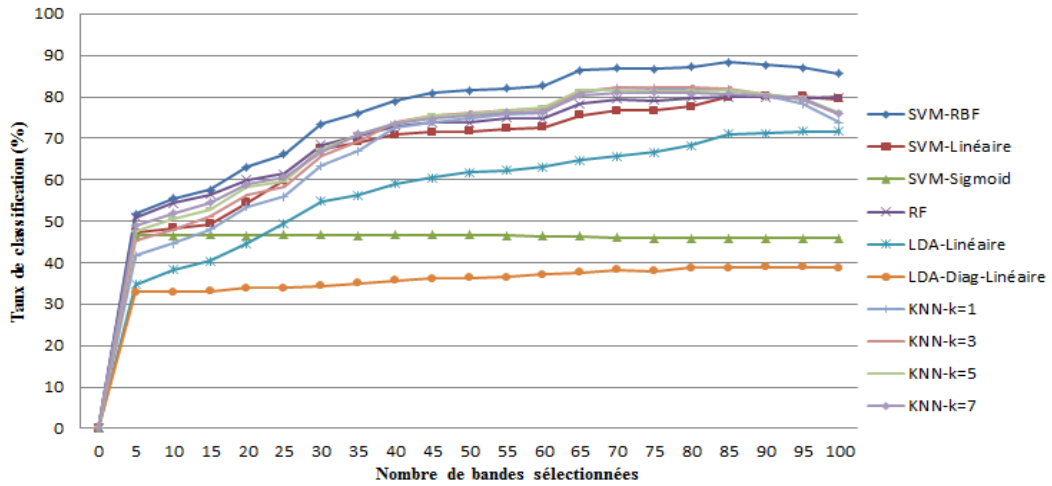


Figure 3. 6: Comparaison des performances en termes du taux de classification et des bandes sélectionnées pour l'image Indian-Pines

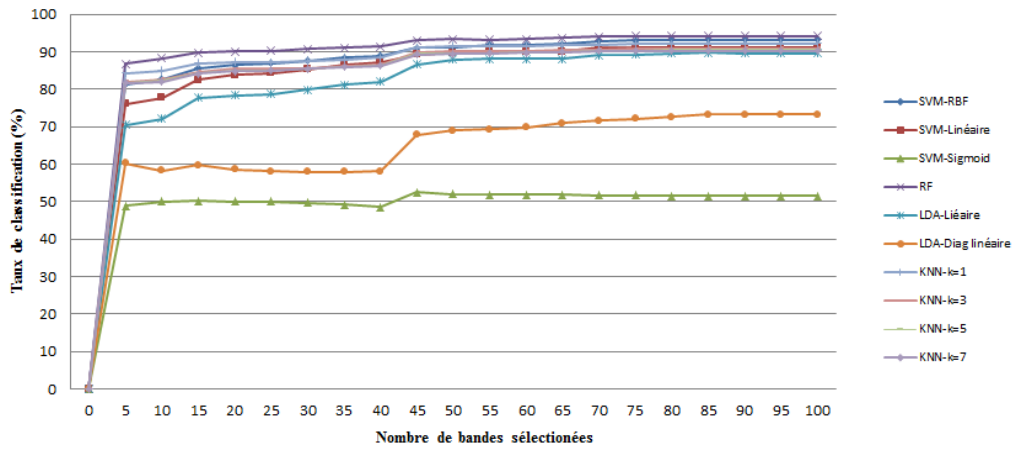


Figure 3. 7: Comparaison des performances en termes de précision de la classification et des bandes sélectionnées pour la scène Salinas.

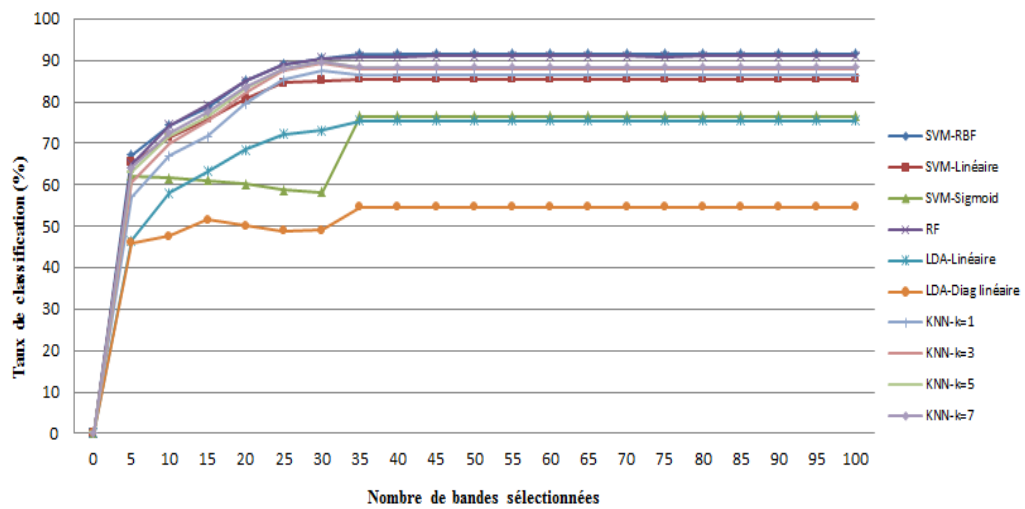


Figure 3. 8: Comparaison des performances en termes de précision de la classification et des bandes sélectionnées pour Pavia-University.

Les résultats obtenus permettent de constater que :

- Les algorithmes de classification étudiés donnent un taux de classification qui atteint 50% lorsque le nombre de bandes sélectionnées est 5 bandes.
- Avec un nombre de bandes sélectionnées proches de 50, le taux de classification s'améliore pour les classifieurs SVM-RBF, SVM-Linear, RF, LDA-Linear, KNN-1, KNN-3, KNN-5 et KNN-7. Les résultats des deux classifieurs SVM-Sigmoid et LDA-Diaglinear restent pratiquement inchangés.
- Les résultats du classifieur SVM avec noyau RBF sont supérieurs aux autres algorithmes en termes du taux de classification. Pour 80 bandes sélectionnées, le taux de classification atteint 87.28% dans le cas de l'IHS Indian Pines (Tableau 3.1), 93.2% pour Salinas (Tableau 3.2) et 91.65% dans le cas de Pavia-University (Tableau 3.3).
- Le classifieur LDA-linéaire produit une meilleure précision que le noyau sigmoïde de SVM et LDA diag-linéaire pour les trois IHS.
- Le taux de classification montre un léger changement au-delà des 80 bandes sélectionnées pour Indian-Pines (Figure 3.6) et Salines (Figure 3.7) vu que les bandes les plus informatives sont déjà sélectionnées. Pour Pavia-University, le taux de classification est pratiquement constant à partir de 35 bandes sélectionnées : elle contient seulement 103 bandes (figure 3.8).
- Pour un nombre de bandes élevé, par exemple 85 bandes dans le cas de Indian-Pines, le taux de classification commence à diminuer, Figure. 3.6 ce qui est équivalent au phénomène de Hughes [Hughes, 1968].

3.4.3.2 Comparaison des algorithmes selon les métriques : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées

Dans cette section, nous comparons les performances des différents algorithmes étudiés selon les paramètres d'évaluation suivants : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution. 80 bandes sont retenues pour cette comparaison. Les résultats sont donnés dans les tableaux 3.4 ; 3.5 et 3.6.

Tableau 3. 4: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas d'Indian-Pines.

		sensibilité (%)	spécificité (%)	Précision (%)	OA (%)	Kappa	Time(s)
SVM	RBF	93.06	99.51	94.42	93.27	0.9282	34.17
	Linéaire	80.35	98.44	85.39	79.01	0.7761	37.80
	Sigmoid	22.66	95.75	20.01	45.72	0.4210	50.34
RF		85.01	99.25	91.68	89.78	0.8910	48.44
DA	Linéaire	77.86	97.87	69.55	69.49	0.6745	2.86
	Diag-linéaire	44.79	95.76	36.58	38.90	0.3483	1.13
KNN	K=1	90.32	99.35	91.33	91.05	0.9045	78.37
	K=3	84.78	99.09	89.43	87.52	0.8669	80.55
	K=5	80.87	98.93	87.64	85.27	0.8428	81.89
	K=7	77.90	98.81	85.82	83.68	0.8260	84.53

Tableau 3. 5: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas de Salinas

		sensibilité (%)	spécificité (%)	Précision (%)	OA (%)	Kappa	Time(s)
SVM	RBF	97.12	99.50	97.41	93.44	0.9300	273.28
	Linéaire	95.68	99.34	96.15	91.30	0.9072	354.14
	Sigmoid	31.74	96.50	27.95	51.52	0.4829	1109.70
RF		98.56	99.78	98.60	97.09	0.9690	256.86
DA	Linéaire	93.87	99.23	93.26	89.56	0.8886	18.04
	Diag-linéaire	78.05	98.09	69.22	72.68	0.7085	5.15
KNN	K=1	98.28	99.71	98.24	96.07	0.9580	2032.16
	K=3	96.48	99.44	96.41	92.50	0.9200	2190.36
	K=5	96.00	99.39	95.99	91.89	0.9135	2485.35
	K=7	95.56	99.35	95.50	91.28	0.9070	2519.85

Tableau 3. 6: Comparaison des taux de classification des quatre algorithmes : SVM, RF, LDA et KNN selon les métriques d'évaluation : sensibilité, spécificité, précision, OA, Kappa et temps d'exécution pour 80 bandes sélectionnées dans le cas de Pavia-University.

		sensibilité (%)	spécificité (%)	Précision (%)	OA (%)	Kappa	Time(s)
SVM	RBF	88.59	98.84	90.07	91.91	0.9090	262.52
	Linéaire	75.17	97.80	68.74	85.48	0.8367	291.08
	Sigmoid	64.78	96.36	59.46	76.38	0.7343	529.15
RF		93.83	99.32	95.35	95.50	0.9493	269.33
DA	Linéaire	78.92	96.84	73.62	75.47	0.7241	8.12
	Diag-linéaire	65.46	94.16	60.42	54.76	0.4911	3.5
KNN	K=1	91.97	99.01	91.75	93.19	0.9234	3435.95
	K=3	88.41	98.61	89.51	90.79	0.8964	3521.30
	K=5	86.98	98.46	88.75	89.87	0.8860	3714.08
	K=7	86.40	98.39	88.62	89.51	0.8819	3886.08

En se basant sur l'évaluation quantitative des résultats présentés dans ces trois tableaux, les remarques suivantes peuvent être formulées :

- Pour l'image hyperspectrale Indian-Pines, les meilleures performances de classification sont obtenues par le classifieur SVM-RBF pour 80 bandes sélectionnées où on obtient 93,06% de sensibilité, 99,51% de spécificité, 94,42% de précision, 93,27% de classification globale OA et un coefficient kappa de 0,9282%. Pour le RF et KNN-1, ils fournissent également de bonnes performances comme on peut le voir dans le tableau 3.4 où par exemple le coefficient kappa est respectivement de 0,8910% et 0,9045% et la précision est respectivement de 91,68% et 91,33%. Le classifieur LDA à noyaux linéaires et diag-linéaires donne des résultats inférieurs à ceux des autres méthodes, Tableau 3.4.
- Pour la scène de Salinas, Tableau 3.5, les résultats indiquent la capacité importante du SVM-RBF et du RF pour séparer les différentes classes par rapport aux autres méthodes. Pour la métrique OA, les algorithmes LDA-diag-linéaire (OA=72.68) et SVM-sigmoïde (OA=51.52%) sont moins performants que le SVM-RBF (OA=93,44%) et le SVM-linéaire (OA =91,30%). Le KNN donne aussi de bons résultats, surtout pour k=1 mais qui sont inférieurs à ceux des algorithmes RF et SVM-RBF. Voir tableau 3.5.

- Dans le cas de l'image hyperspectrale Pavia-University, le tableau 3.6 résume les résultats obtenus en utilisant les différents algorithmes selon les différentes métriques. Les résultats montrent que les classifieurs SVM-RBF et RF fournissent de meilleures valeurs des métriques par rapport au KNN, avec les différentes valeurs de k, et le LDA.
- En termes du temps d'exécution, nous remarquons que :
 - Premièrement, le temps de simulation dépend de la taille des images et de l'algorithme d'apprentissage comme on peut le constater pour les trois IHS Tableaux 3.4, 3.5 et 3.6. L'intervalle du temps d'exécution varie entre 1.13s et 84.53s pour Indian-Pines (Tableau 3.4). Pour Salinas, cette plage est plus élevée puisque la taille de l'image augmente, le temps est entre 5,15s et 2519,9s (tableau 3.5). Pour Pavia-University, il nous faut encore plus de temps de 3,5s à 3886,1s pour séparer toutes les classes (Tableau 3.6).
 - Deuxièmement, les classifieurs SVM-RBF et RF fournissent de bonnes performances de classification avec un avantage en termes de temps de fonctionnement. Le KNN donne aussi de bons résultats mais il nécessite un temps d'exécution qui augmente avec le nombre des plus proches voisins utilisés (la valeur de k). Le LDA, par contre, est le plus rapide, mais il est moins performant.

3.4.3.3 Visualisation des scènes classifiées

La figure 3.9 (respectivement 3.10 et 3.11) présente la vérité de terrain d'Indian-Pines (respectivement Salinas et Pavia-University) ainsi que leurs reproductions par les classifieurs suivants : SVM-RBF, RF, KNN-1 et LDA-linear. Par l'inspection visuelle, nous constatons que le SVM-RBF donne la carte classifiée la plus proche de la vérité de terrain : nous avons une discrimination notable de toutes les classes. Le RF et le KNN-1 reproduisent mieux la vérité de terrain de Salinas que le SVM-RBF mais ils demandent plus de temps d'exécution. La carte la moins illustrative est obtenue en utilisant le classifieur LDA-linear pour les trois scènes.

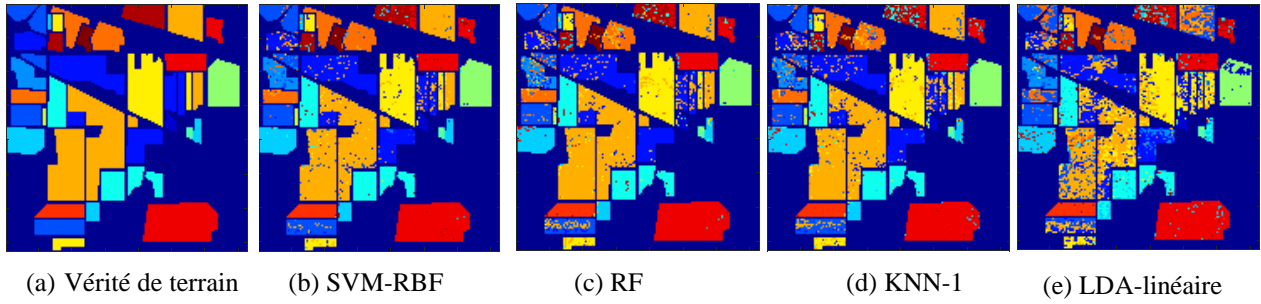


Figure 3. 9: La vérité de terrain de Indian-Pines (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.

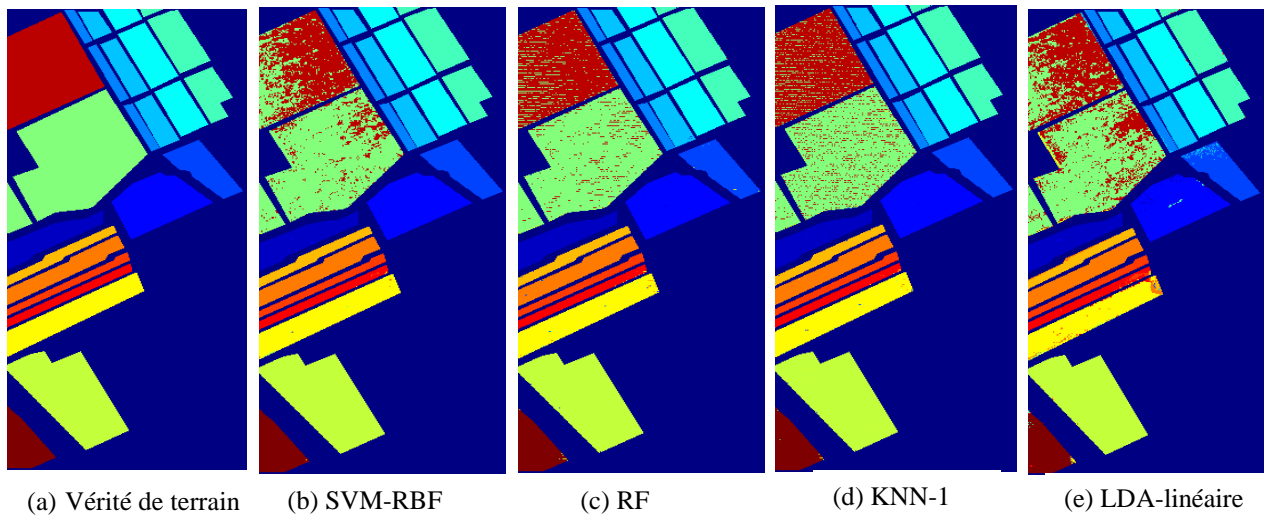


Figure 3. 10: La vérité de terrain de Salinas (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.

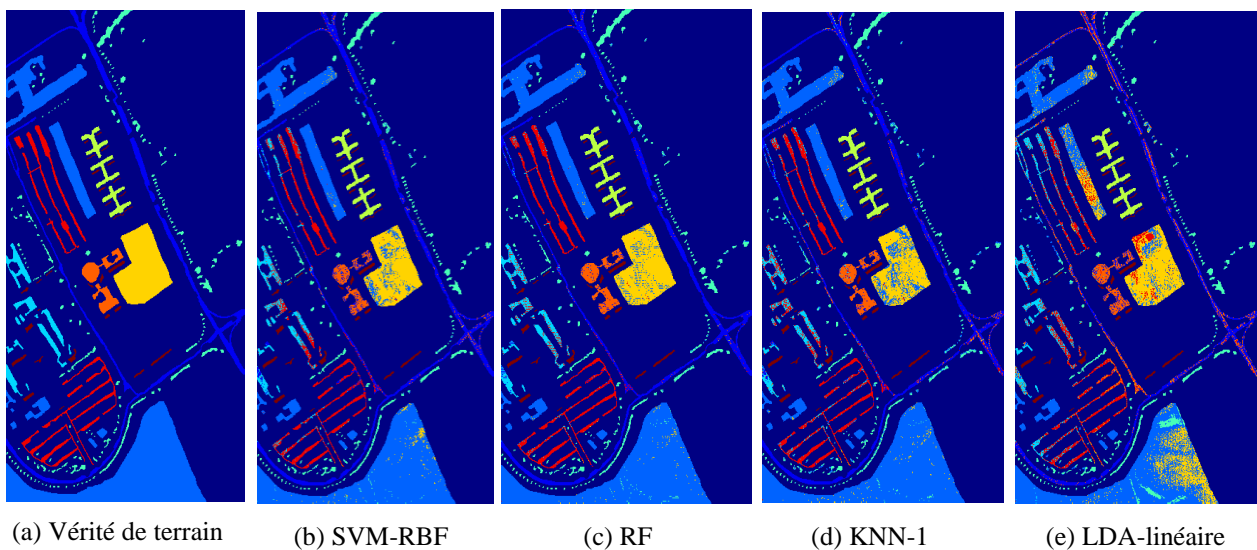


Figure 3. 11: La vérité de terrain de Pavia (a) et sa reproduction en utilisant SVM-RBF (b), RF (c), KNN-1 (d) et LDA-Linéaire (e), pour 80 bandes sélectionnées.

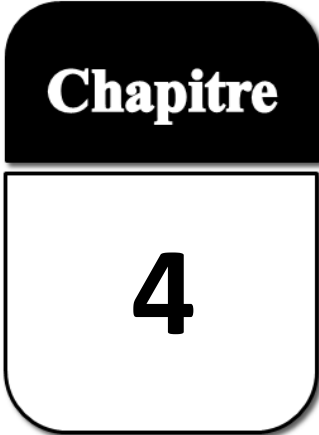
3.5 Conclusion

Dans ce chapitre, nous avons évalué les performances de quatre algorithmes de classification appliqués aux images hyperspectrales. Pour choisir les bandes les plus informatives et faire face au phénomène de Hughes, la réduction de dimensionnalité de ces IHS a été réalisée en utilisant l'information mutuelle. Les classifieurs étudiés sont : Les Machines à Vecteurs de Support SVM avec les noyaux (RBF, Linear et Sigmoid), Les forêts aléatoires RF, l'Analyse discriminante linéaire LDA avec les noyaux (Linear et Diag-linear), et K-plus proches voisins KNN avec (k=1, 3, 5 et 7).

Les algorithmes ont été évalués à l'aide de trois images hyperspectrales de télédétection collectées par les capteurs hyperspectraux AVIRIS de la NASA et ROSIS. Ces images présentaient des caractéristiques différentes en termes de nombre et type des classes et aussi de résolutions spectrales et spatiales.

Pour l'évaluation des performances de chaque classifieur, plusieurs paramètres ont été calculés : sensibilité, spécificité, précision, précision globale OA, coefficient kappa k et temps de calcul. Ces métriques affirment que le SVM-RBF fournit les performances maximales suivi par RF. En termes de compromis entre les performances de classification et le temps d'exécution, le SVM-RBF est meilleur que le RF : le SVM-RBF nécessite un temps réduit pour séparer efficacement les différentes classes des scènes hyperspectrales. Le KNN donne aussi de bonnes performances de classification, mais il nécessite un temps largement considérable. Le LDA, même s'il est le plus rapide, il reste le moins performant en comparaison avec tous les algorithmes étudiés.

Ce travail nous a permis de retenir le SVM-RBF parmi les classifieurs étudiés (SVM-RBF, SVM-Linear, SVM-Sigmoid, RF, LDA-Linear, LDA-Diaglinear, KNN-1, KNN-3, KNN-5, KNN-7) pour la validation des algorithmes de réduction de dimensionnalité que nous proposons dans le reste de ce mémoire.



Chapitre
4

Une nouvelle approche filtre pour la sélection de bandes et la classification des images hyperspectrales basée sur l'information mutuelle normalisée et les machines à vecteurs de support

Résumé :

Ce chapitre introduit une nouvelle approche filtre pour la réduction de la dimensionnalité et la classification des images hyperspectrales en utilisant la théorie de l'information (information mutuelle normalisée) et les machines à vecteurs de support SVM. Cette méthode consiste à sélectionner un sous-ensemble minimal de bandes les plus informatives et les plus pertinentes parmi les données d'entrée pour une meilleure efficacité de classification. Nous avons appliqué notre algorithme proposé sur les deux images d'AVIRIS « Indian Pines » et « Salinas ». Les résultats expérimentaux ont été évalués par différentes métriques largement utilisées dans ce domaine. La comparaison avec d'autres méthodes dans la littérature prouve que notre méthode peut produire de bonnes performances avec un nombre réduit de bandes sélectionnées dans un bon intervalle de temps d'exécution.

4.1 Introduction

La sélection des bandes est une tâche ardue dans la classification des images hyperspectrales de télédétection IHS. Cela résulte de sa haute résolution spectrale, ses multi-classes et du nombre limité d'échantillons d'apprentissage qui mène au problème de la malédiction de dimensionnalité ou phénomène de Hughes [Hughes, 1968]. Ce chapitre propose une nouvelle approche filtre basée sur l'Information mutuelle Normalisé (NMI) pour la sélection du sous-ensemble optimal de bandes pertinentes et non redondantes.

La suite de ce chapitre est organisée comme suit: la section suivante présente les travaux relatifs à la sélection des attributs pour la réduction de la dimensionnalité des images hyperspectrales. Dans la section 3, nous décrivons l'approche proposée. La section 4 présente les bases de données et discute les résultats expérimentaux en comparaison avec d'autres méthodes récentes [Peng, 2005 ; Medjahed, 2016].

4.2 Les recherches antérieures

Les méthodes de sélection des attributs suscitent un intérêt croissant au niveau de la recherche scientifique au cours des dernières décennies. Avec l'acquisition des données d'énorme quantité d'information et les défis qui en résultent, de nouveaux problèmes apparaissent et de nouvelles approches pour la sélection des attributs sont toujours en demande pour surmonter les défis de la classification des IHS. La méthode de sélection des attributs basée sur des critères d'information mutuelle de max-pertinence et min-redondance MRMR_[Peng, 2005] est utilisée pour sélectionner les attributs pour la classification des IHS. Guo [Guo, 2008] présente une stratégie d'optimisation « greedy » pour sélectionner des bandes à partir des IHS. Ren [Ren, 2011] a proposé un regroupement adaptatif pour la sélection des bandes. Zhu [Zhu, 2016] propose une méthode non supervisée pour la sélection de bandes par extraction de groupe dominant DSEBS en utilisant la corrélation structurelle.

De nouvelles méthodes apparaissent encore dans la littérature. Medjahed [Medjahed, 2016] utilise « Gray Wolf Optimizer GWO » pour reformuler la sélection d'attributs en tant que problème combinatoire basé sur la séparabilité des classes et le taux de précision en modélisant cinq fonctions objectives. D'autre part, Wu [Wu, 2018] utilise une analyse discriminante de Fisher locale semi-supervisée basée sur des échantillons de pseudo-étiquettes pour la réduction de la dimensionnalité des IHS. Dans notre travail, nous proposons une nouvelle approche filtre nommée NMIBS basée sur la théorie de l'information. Nous utilisons

l'information mutuelle normalisée NMI avec les Machines à Vecteurs de Support SVM pour la réduction de dimensionnalité et la classification des IHS [Nhaila, 2018b].

Pour confirmer l'efficacité de l'approche proposée, les expérimentations ont été effectuées sur les deux images hyperspectrales réelles d'AVIRIS de la NASA: Indian Pines et Salinas. La comparaison est faite avec plusieurs techniques de sélection de bandes et de classification des images hyperspectrales.

4.3 Méthodologie proposée

4.3.1 Principe

L'objectif principal de ce travail est d'améliorer les performances de classification des images hyperspectrales par information mutuelle en introduisant une nouvelle approche filtre pour la sélection de bandes. Elle consiste à sélectionner le sous-ensemble optimal de bandes pertinentes et à éliminer les bandes bruyantes et redondantes à l'aide de l'information mutuelle normalisée NMI.

Selon le principe général des méthodes de sélection d'attributs [Dash, 1997], l'organigramme de notre méthode proposée peut être schématisé comme suit :

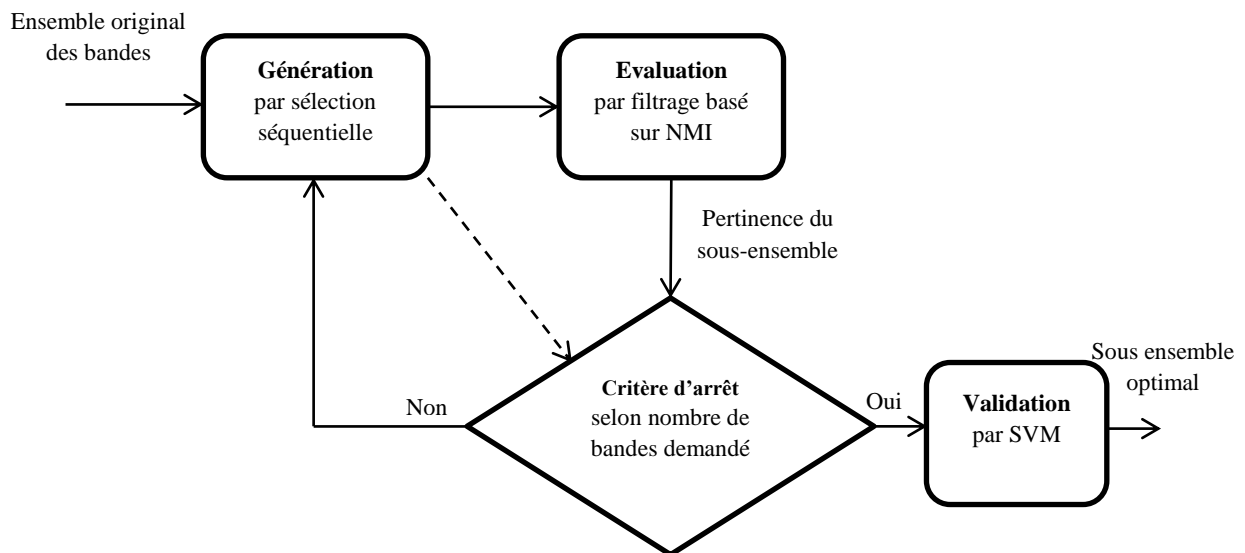


Figure 4. 1: Organigramme de la méthode de sélection avec validation proposée.

Notre algorithme comprend quatre étapes principales :

- La procédure de génération des bandes candidates en utilisant la sélection séquentielle de bandes en commençant par un ensemble vide.

- La fonction d'évaluation qui permet de juger la qualité du sous-ensemble actuel. Dans cette étape, nous mesurons les informations des bandes et leur dépendance en utilisant l'information mutuelle normalisée NMI.
- Le critère d'arrêt pour décider d'arrêter la recherche. Cela dépend du nombre d'itérations et des bandes à sélectionner.
- La procédure de validation pour tester l'efficacité du sous-ensemble de bandes retenu. Au cours de cette étape, nous avons appliqué le classifieur SVM à des données de référence du monde réel. Les résultats obtenus ont été comparés avec les méthodes les plus récentes.

Le reste de cette section traite la sélection de bande à l'aide de l'information mutuelle normalisée et présente le processus de sélection complet de l'algorithme proposé.

4.3.2 Sélection des bandes en utilisant l'information mutuelle normalisée

4.3.2.1 Formulation de l'information mutuelle normalisée NMI

L'information mutuelle IM a été utilisée pour mesurer la similarité entre les bandes et la vérité de terrain. L'IM prend des valeurs élevées en cas de grande similarité ce qui a permis de sélectionner des sous-ensembles réduits de bandes les plus informatives. Cependant, plusieurs chercheurs surtout au domaine médical ont démontré que l'utilisation de la forme normalisée de l'information mutuelle est plus efficace pour mesurer la similarité entre les images. Ainsi plusieurs formules ont été proposées dans la littérature pour la normalisation de l'information mutuelle.

Estévez [Estévez, 2009] propose d'utiliser la moyenne de l'information mutuelle normalisée comme mesure de la redondance entre les attributs. Il normalise l'IM entre X et Y par l'entropie minimale des deux attributs. Elle est définie par :

$$NMI(X, Y) = \frac{IM(X, Y)}{\min\{H(X), H(Y)\}} \quad (4.1)$$

Studholme [Studholme, 1999] introduit une mesure normalisée de l'IM qui représente le rapport entre la somme des entropies marginales et l'entropie conjointe de X et Y, Formule 4.2. Les résultats de ces expériences sur des images médicales 3D, ont montré que la formule proposée améliore considérablement la robustesse de l'alignement des images étudiées. Cette mesure est définie comme :

$$NMI(X, Y) = \frac{H(X)+H(Y)}{H(X,Y)} \quad (4.2)$$

Une autre formule de normalisation de l'IM a été introduite par Strehl [Strehl, 2002] pour faciliter l'interprétation des valeurs qui seront bornées entre 0 à 1. C'est une normalisation par la moyenne géométrique de H(X) et H(Y). Elle est basée sur l'observation que $I(X ; Y)$ est inférieur au $\min(H(X);H(Y))$ et que $H(X) = I(X;X)$. Cette formule a été utilisée dans son travail [Strehl, 2002] pour résoudre le problème de combiner plusieurs partitions d'un ensemble d'objets en un seul cluster consolidé sans accéder aux caractéristiques ou algorithmes qui déterminent ces partitions. Cette formule est définie par :

$$NMI(X, Y) = \frac{IM(X,Y)}{\sqrt{H(X).H(Y)}} \quad (4.3)$$

D'autres formes normalisées ont été proposées dans la littérature [Kvalseth , 1987 ; Oribe, 1985] comme suit:

$$NMI(X, Y) = \frac{IM(X,Y)}{\frac{1}{2}(H(X)+H(Y))} = 2 \frac{IM(X,Y)}{H(X)+H(Y)} \quad (4.4)$$

$$NMI(X, Y) = \frac{IM(X,Y)}{H(X)} \quad (4.5)$$

$$NMI(Y, X) = \frac{IM(X,Y)}{H(Y)} \quad (4.6)$$

Dans ce travail, nous utiliserons l'information mutuelle normalisée proposée par [Studholme, 1999] et définie comme suit:

$$NMI(GT, B) = \frac{H(GT)+H(B)}{H(GT,B)} \quad (4.7)$$

Cette mesure représente le rapport entre la somme de l'information totale (contenue dans la vérité du terrain GT et la bande candidate B) et l'information conjointe (de GT et B).

4.3.2.2 Sélection des bandes pertinentes

Rappelons que la formule de l'information mutuelle, équation (3.4), est la différence entre l'information totale contenue dans deux images (GT et B) et leur information conjointe:

$$MI(GT, B) = H(GT) + H(B) - H(GT, B) \quad (4.8)$$

Ainsi, maximiser l'information mutuelle revient à maximiser l'information totale des deux images (entropies) tout en minimisant l'information conjointe (entropie conjointe).

La formule de normalisation retenue consiste à diviser par l'entropie conjointe $H(GT, B)$.

Rappelons que :

$$H(GT, B) \leq H(GT) + H(B) \quad (4.9)$$

Avec égalité si et seulement si GT et B sont indépendants. Ainsi, l'information mutuelle normalisée satisfait :

$$NMI(GT, B) = \frac{H(GT)+H(B)}{H(GT,B)} \geq 1 \quad (4.10)$$

Considérant les deux IHS d'AVIRIS Indian-Pines et Salinas, Pour chaque IHS, nous allons estimer l'interdépendance entre ses bandes spectrales et la vérité de terrain correspondante. Nous calculons l'information mutuelle normalisée NMI et l'information mutuelle IM en utilisant respectivement les équations 4.7 et 4.8. Les résultats sont présentés à la Figure 4.2 (à gauche) dans le cas d'Indian-Pines et (à droite) dans le cas de Salinas.

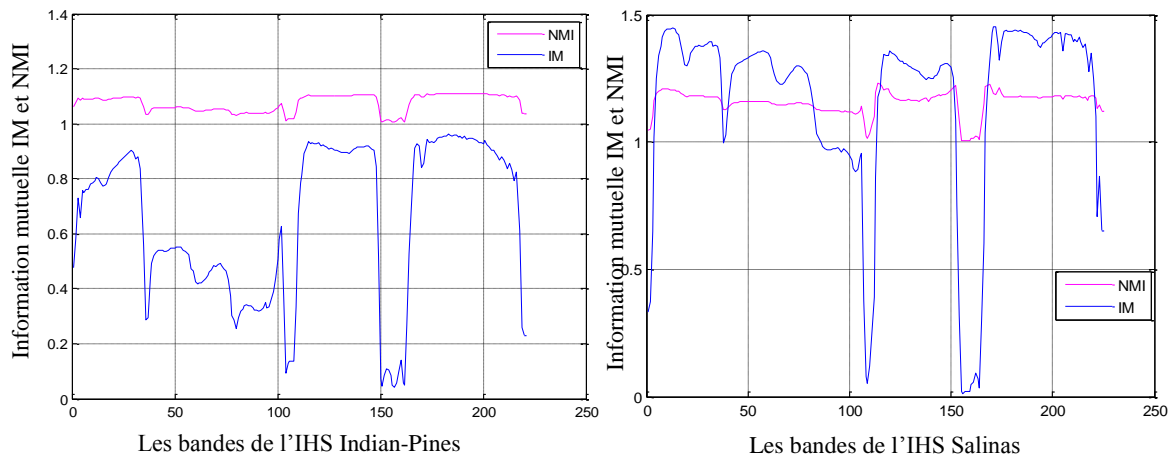


Figure 4. 2: L'Information mutuelle normalisée NMI et l'Information mutuelle IM des bandes spectrales d'Indian-Pines (à gauche) et Salinas (à droite).avec les GT correspondantes.

La figure montre que les bandes ont différentes valeurs d'IM et de NMI avec la vérité de terrain selon les zones spectrales. Ces valeurs correspondent au niveau de similarité entre les bandes et les vérités de terrain : cette similarité est élevée, moyenne où faible.

Prenons l'exemple de sélection d'un sous ensemble de cinq bandes qui ont les plus grandes valeurs d'IM et de NMI avec la vérité de terrain dans le cas d'Indian-Pines.

- En utilisant l'IM, le sous-ensemble sélectionné contient les bandes suivantes:
 $S_{IM} = \{B183, B184, B186, B187, B191\}$.

- En utilisant la NMI, le sous-ensemble des bandes sélectionnées contient bandes suivantes: $S_{NMI} = \{B191, B186, B189, B201, B187\}$.

Nous remarquons que les bandes sélectionnées en utilisant les deux formules sont différentes.

En sélectionnant maintenant un sous ensemble de 10 bandes.

- En utilisant l'IM, le sous-ensemble sélectionné contient les bandes suivantes: $S_{IM} = \{B183, B184, B186, B187, B191, B189, B182, B185, B181, B179\}$.
- En utilisant la NMI, le sous ensemble des bandes sélectionnées contient bandes suivantes:
 $S_{NMI} = \{B191, B186, B189, B201, B187, B183, B184, B194, B195, B190\}$.

En utilisant ces sous-ensembles de bandes sélectionnées, nous calculons la précision globale de classification OA et la précision moyenne de classification de toutes les classes AA. Les résultats obtenus sont présentés dans le tableau suivant.

Tableau 4. 1: Précision globale de classification OA et précision moyenne de classification AA selon le groupe de bandes sélectionnées en utilisant l'IM et la NMI pour Indian-Pines.

	OA(%)		AA(%)	
	S=5	S=10	S=5	S=10
IM	51.52	55.91	30.80	40.78
NMI	52.99	57.33	36.21	43.36

D'après ces illustrations, on constate que la sélection des bandes en utilisant la NMI est plus performante qu'en utilisant l'IM. Ainsi le critère de sélection basé sur la NMI stipule que la bande doit être choisie lorsque sa valeur de NMI est élevée. Ceci signifie qu'une grande quantité d'information peut être obtenue en considérant cette bande. Ce constat va nous permettre de modéliser la pertinence d'une bande pour la classification et sélectionner ainsi les bandes informatives et éliminer celles qui sont bruyantes. La NMI sera utilisée dans les étapes de génération et d'évaluation de notre méthodologie proposée comme indiqué dans l'algorithme proposé dans la section suivante.

4.3.3 Algorithme proposé

Le processus de sélection complet de la méthodologie proposée est le suivant:

Algorithme

Entrée: Image hyperspectrale

GT: Vérité de terrain;

B:={Bandes};

Sortie:

Le sous ensemble des bandes sélectionnées.

Carte Classifiée.

1 (Initialisation)

$B \leftarrow$ les bandes de l'IHS

$S \leftarrow \{\}$ // le groupe des bandes sélectionnées

$L \leftarrow$ Nombre des bandes à sélectionner

$z \leftarrow$ Nombre des itérations

$N \leftarrow$ Nombre des bandes de l'IHS

$Th \leftarrow$ Seuil de contrôle de redondance

2) Pour toutes B_i dans B , Calculer l'information mutuelle normalisée NMI avec la vérité de terrain en utilisant l'équation 4.7:

Pour $b_i \in B$ calculer $NMIBands = NMI(GT, b_i)$

3) Processus de sélection

Sélectionner la première bande b_i qui maximise $NMIBands$:

$i := \operatorname{argmax}(NMI(B))$

$B \leftarrow B \setminus \{b_i\}; S \leftarrow \{b_i\};$

$GT_{est0} =$ bande b_i ; calculer $VNMI^* = NMI(GT, GT_{est0})$;

$Z = 0$;

Tant que $|S| < L$ & $z < N-1$ Faire

$i := \operatorname{argmax} NMI(B \setminus S) \ \& \ B \leftarrow B \setminus \{b_i\}$

$GT_{est} := (GT_{est0} + \text{band}(b_i)) / 2$

Calculer $VNMI = NMI(GT, GT_{est})$

Si $VNMI > VNMI^* + Th$ alors

$VNMI^* := VNMI$;

$GT_{est0} := GT_{est}$;

$S \leftarrow S \cup \{b_i\};$

Fin Si

$z \leftarrow z + 1$

Fin Tant que

4) S contient les bandes sélectionnées

5) Carte classifiée = SVM(S)

Dans cet algorithme, pour générer le sous-ensemble optimal de bandes réduites, nous initialisons les bandes sélectionnées par celle qui a la plus grande NMI avec la vérité du

terrain, puis une carte de référence approchée appelée GT_{est} est construite par la moyenne de la dernière initialement nommée GT_{est0} avec la première bande. Chaque bande candidate est conservée si elle augmente la dernière valeur de $NMI(GT, GT_{est})$ utilisée comme fonction d'évaluation, sinon, elle sera rejetée. Le seuil Th est introduit pour contrôler la redondance autorisée, la valeur de Th est déterminée d'une manière empirique. Le critère d'arrêt est testé en fonction du nombre de bandes à sélectionner k et du nombre total d'itérations. Enfin, l'étape de validation est réalisée à l'aide du classifieur SVM pour produire les cartes classifiées en utilisant les bandes sélectionnées par cet algorithme. Plusieurs mesures d'évaluation sont ensuite calculées à partir de la matrice de confusion pour la comparaison avec d'autres méthodes.

4.4 Résultats expérimentaux et discussion

4.4.1 Bases de données utilisées

Afin d'évaluer les performances de l'approche proposée, les expériences sont conduites sur les deux images hyperspectrales d'AVIRIS : Indian Pines et Salinas. Le Tableau 4.2 rappelle les principales caractéristiques de ces images. L'image couleur composite et la vérité de terrain correspondante avec les classes sont présentées la Figure 4.3 en (a) pour Indian Pines et en (b) pour Salinas.

Tableau 4. 2: *Caractéristiques des images hyperspectrales utilisées dans les expériences.*

	Nombre de bandes	Nombre de classes	Taille des images	Région des longueurs d'ondes	Résolution Spatiale
Indian	224	16	145x145	0.4-2.5 μm	20m pixels
Salinas	224	16	217x512	0.4-2.5 μm	3.7m pixels

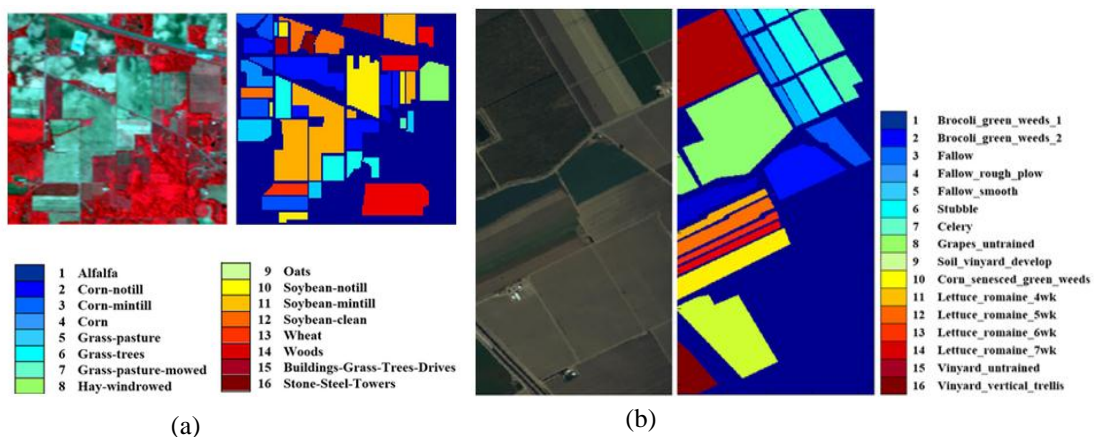


Figure 4. 3: *L'image couleur composite et la vérité de terrain correspondante avec les classes pour: les IHS (a) Indian Pines et (b) Salinas.*

4.4.2 Protocole expérimental

Le SVM-RBF est retenu dans ce travail suite à l'étude effectuée au chapitre précédent. Il est très adapté à la nature des images hyperspectrales et a été appliqué avec succès dans plusieurs applications des IHS dans le monde réel [Nhaila, 2018a; Xie, 2017]. Trois ensembles d'apprentissages (10%, 25% et 50%) sont construits d'une manière aléatoire pour entraîner le classifieur.

La méthode proposée est comparée à diverses méthodes de sélection d'attributs, notamment MIM [Viola, 1997], MRMR [Peng, 2005] et GWO [Medjahed, 2016]. La classification par SVM en utilisant toutes les bandes sans réduction de dimensionnalité est également incluse dans la comparaison. Les étapes des algorithmes MRMR et GWO sont présentées en annexe.

Toutes les expériences sont compilées sur un ordinateur quadri-cœur Duo, 64-b, à fréquence de CPU 2.1 Ghz avec 3 Go de RAM. L'algorithme proposé s'arrête lorsque le nombre désiré de bandes sélectionnées est atteint.

4.4.3 Résultats et discussion

Les résultats expérimentaux sur les images hyperspectrales Indian Pines et Salinas en utilisant l'approche proposée sont présentés dans cette section. Les performances de la classification sont évaluées à l'aide de deux mesures d'évaluation largement utilisées dans les applications des images hyperspectrales en télédétection, à savoir: la précision moyenne « Average Accuracy » AA et la précision globale « Overall Accuracy » OA. La AA mesure la moyenne de l'exactitude de la classification pour toutes les classes; elle est calculée comme étant le rapport de la somme de la précision individuelle de chaque classe sur le nombre des classes. Tandis que l'OA indique le nombre de pixels correctement prédits sur tous les échantillons de test. Le temps de calcul est également calculé.

Les Tableaux 4.3 et 4.4 montrent respectivement la précision globale obtenue pour les IHS Indian-Pines et Salinas. La première colonne de chaque tableau représente le nombre de bandes sélectionnées. Les colonnes restantes indiquent les résultats obtenus en utilisant différents pourcentages d'échantillons d'entraînement (10%, 25% et 50%).

Tableau 4. 3: La précision globale OA Obtenue en utilisant l'algorithme proposé sur l'image hyperspectrale Indian-Pines pour différents nombre de bandes sélectionnées et différents pourcentages d'échantillons d'apprentissage.

Nombre de bandes sélectionnées	Ensemble d'apprentissage 10%	Ensemble d'apprentissage 25%	Ensemble d'apprentissage 50%
10	55.2	56.72	57.33
20	59.65	61.28	62.76
30	68.23	71.61	73.98
40	72.06	77.29	81.93
50	74.00	79.65	84.84
60	76.41	83.63	88.63
70	77.60	84.56	90.24
80	77.83	84.55	90.74
90	80.90	86.98	93.90
100	80.83	87.25	93.48
Toutes les bandes	60.74	69.42	75.72

Tableau 4. 4: La précision globale OA Obtenue en utilisant l'algorithme proposé sur l'image hyperspectrale Salinas pour différents nombre de bandes sélectionnées et de pourcentage d'échantillons d'apprentissage.

Nombre de bandes sélectionnées	Ensemble d'apprentissage 10%	Ensemble d'apprentissage 25%	Ensemble d'apprentissage 50%
10	80.35	81.36	81.81
20	88.13	88.54	88.90
30	89.84	90.27	90.58
40	90.66	91.29	91.63
50	91.80	92.41	92.79
60	92.26	92.86	93.27
70	92.59	93.23	93.59
80	92.65	93.28	93.80
90	92.62	93.36	93.91
100	92.65	93.48	94.08
Toutes les bandes	87.31	88.77	90.02

À partir de ces résultats, on constate que:

- Premièrement, il est évident que le nombre de pixels utilisés pour l'entraînement influe sur le taux de classification. L'OA augmente avec la taille des ensembles d'entraînement utilisés dans les deux images Indian et Salinas. Par exemple, avec 70 bandes sélectionnées dans la scène Indian-Pines, nous obtenons 77,60%, 84,55% et 90,24% pour respectivement 10%, 25% et 50% des données d'entraînements, comme indiqué dans le tableau 4.3. Pour Salinas, nous obtenons respectivement 92,59%, 93,23% et 93,59%, voir tableau 4.4. Les cartes classifiées obtenues pour ces valeurs sont illustrées à la Figure 4.4 pour Indian-Pines et à la Figure 4.5 pour la scène Salinas.

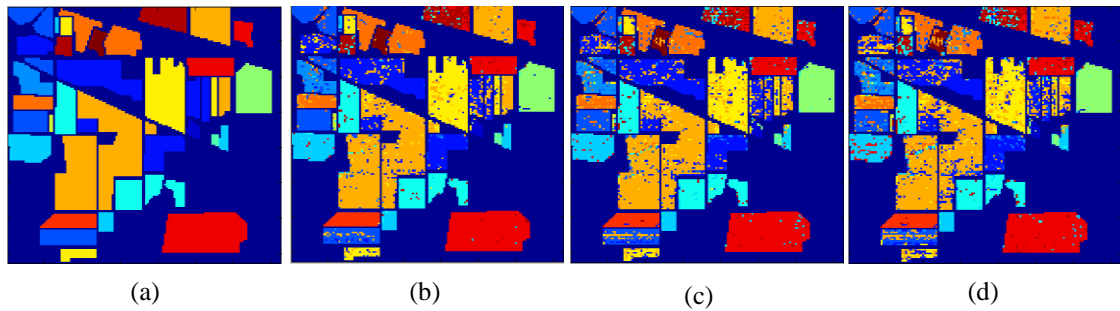


Figure 4. 4: La vérité-terrain d'Indian-Pines (a) et les images classifiées en utilisant la méthode proposée pour différents pourcentages de données d'entraînement: 50% (b), 25% (c) et 10% en (d).

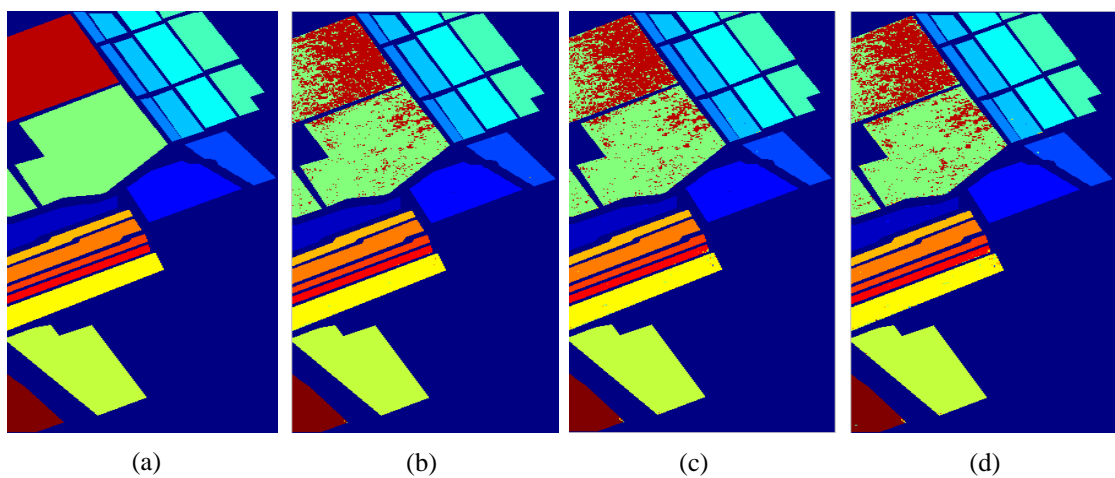


Figure 4. 5: La vérité-terrain de Salinas (a) et les images classifiées en utilisant la méthode proposée pour différents pourcentages de données d'entraînement: 50% (b), 25% (c) et 10% en (d).

- Deuxièmement, la combinaison de l'information mutuelle normalisée et le classifieur SVM dans la méthodologie proposée produit de bons résultats de classification, même avec un nombre limité de pixels d'apprentissage. Par exemple, dans le cas de 10% et avec seulement 40 bandes sélectionnées sur 224, l'OA atteint 72,06% pour Indian-Pines et 90,66% pour Salinas.
- Troisièmement, il est clair que l'utilisation d'un sous-ensemble de bandes pertinentes donne des résultats de classification meilleurs que l'utilisation de toutes les bandes. Dans Indian-Pines, voir le tableau 4.3, la valeur OA obtenue en utilisant toutes les bandes est égale à 75,72%, alors qu'elle atteint 90,24% avec un sous-ensemble réduit de 70 bandes. À Salinas, nous obtenons 87,31% en considérant toutes bandes contre 90,66% pour seulement 40 bandes sélectionnées ce qui confirme l'efficacité de la

méthodologie proposée pour sélectionner un ensemble réduit de bandes optimales et éliminer les bandes redondantes et bruyantes qui réduisent le taux de classification.

Dans les expériences suivantes, l'approche proposée est comparée à d'autres méthodes définies dans la littérature en utilisant seulement 10% de données d'entraînement. Les résultats obtenus sont présentés dans le tableau 4.5 et évalués en utilisant AA, OA et le temps d'exécution.

Tableau 4. 5: La précision moyenne AA (%), la précision globale OA (%) et le temps de calcul (s) obtenus par l'algorithme proposé en comparaison avec différentes méthodes sur les bases de données Indian-Pines pour 75 bandes sélectionnées et Salinas pour 69 bandes sélectionnées

Les méthodes	Indian-Pines			Salinas		
	AA(%)	OA(%)	Temps(s)	AA(%)	OA(%)	Temps(s)
SVM-toutes les bandes	42.67	60.74	42.83	91.45	87.31	397.47
MIM	56.06	73.54	12.05	93.45	88.91	126.24
MRMR	58.70	75.70	24.87	93.56	89.67	151.55
Gwo-J1	67.82	71.28	170.3	94.46	89.07	1166
Gwo-J2	62.57	67.44	1.7	94.68	89.25	1.05
Gwo-J3	64.10	70.29	0.48	94.89	89.41	5.34
Gwo-J4	73.89	73.67	250	97.37	95.38	1221
Gwo-J5	70.43	70.65	197	95.50	90.80	1198
NMIBS proposé	70.41	77.90	8.77	96.47	92.54	84.67


- Les résultats les plus faibles sont obtenus par la classification SVM en utilisant toutes les bandes. Cela confirme l'importance de la réduction de dimensionnalité comme prétraitement de la classification des images hyperspectrales ; ceci dans le but d'éliminer les bandes non pertinentes.
- La méthode MIM offre des performances inférieures à celles des autres méthodes de réduction de dimensionnalité (MRMR, GWO et le NMIBS proposé) sauf pour l'OA obtenu avec les fonctions J1, J2, J3 et J5 de la méthode GWO dans le cas de l'IHS Indian-Pines. Les performances du MIM sont faibles puisqu'elle sélectionne les bandes uniquement en se basant sur la maximisation de l'information mutuelle sans traitement de la redondance entre bandes.
- Gwo-J4 dépasse notre méthode, mais elle nécessite beaucoup plus de temps de (250 secondes contre seulement 8,77 secondes) pour Indian-Pines. Dans le cas de Salinas, le temps d'exécution de Gwo-J4 est de 1121 secondes contre seulement 84,67 secondes.

4.5 Conclusion

Dans ce chapitre, nous avons proposé une nouvelle méthode de sélection de bandes basée sur l'information mutuelle normalisée et le classifieur SVM. L'objectif est réduire l'effet de la malédiction de la dimensionnalité, sélectionner les bandes informatives et enlever celles qui sont redondantes ou bruitées. Cela améliore la classification des images hyperspectrales. Cette approche proposée est de type Filtre. Elle adopte l'information mutuelle normalisée pour générer et évaluer la pertinence des sous-ensembles de bandes sélectionnées. La validation est effectuée à l'aide du classifieur supervisé SVM avec le noyau RBF.

Les expériences sont réalisées sur deux IHS de référence, collectées par le capteur hyperspectral AVIRIS. Les résultats obtenus en termes de taux de classification global OA permettent de conclure que l'effet de sélection par la méthode proposée augmente les performances de séparation des classes dans IHS.

La comparaison avec d'autres méthodes dans la littérature est réalisée selon des paramètres largement utilisés dans ce domaine : AA et OA et le temps d'exécution. Les résultats montrent l'efficacité et la rapidité de notre approche qui est de type Filtre. Cependant, les approches Filtres n'intègrent pas les performances des classifieurs dans l'étape de génération des sous-ensembles de bandes. Dans le chapitre suivant, nous introduisons une stratégie Wrapper qui fait appel au classifieur SVM-RBF durant la sélection des bandes.



Chapitre
5

**Nouvelle méthode Wrapper basée sur
l'information mutuelle normalisée pour la
réduction de dimensionnalité et la classification
des images hyperspectrales**

Résumé :

Dans ce chapitre, nous proposons une nouvelle approche Wrapper basée sur l'information mutuelle normalisée (NMI), la probabilité d'erreur (PE) et les Machines à Vecteurs de Support (SVM) pour réduire la dimensionnalité et la classification des images hyperspectrales. Cela permet d'augmenter l'efficacité de la reproduction des cartes thématiques. Les expériences sont réalisées sur les deux images hyperspectrales d'AVIRIS de la NASA : Indian-Pines et Salinas. Les métriques AA (la moyenne des classifications des classes), OA (la précision globale de classification), Kappa (le coefficient de concordance) et ICA (la classification individuelle de chaque classe) ont été calculées pour évaluer les performances de l'algorithme proposé. Les résultats obtenus prouvent que notre méthode augmente les performances de classification et fournit une carte thématique précise par rapport à d'autres algorithmes. Cette méthode peut être améliorée pour plus d'efficacité de classification.

5.1 Introduction

Ce chapitre introduit un nouvel algorithme Wrapper nommé (WNMIPE) basé sur l'information mutuelle normalisée NMI et la probabilité d'erreur PE avec les machines à vecteurs de support SVM. Le but est la réduction de la dimensionnalité des IHS pour améliorer l'efficacité de la classification. L'évaluation des performances de la méthode proposée est basée sur plusieurs métriques. Les expériences et la comparaison avec d'autres algorithmes sont effectuées en utilisant les deux images hyperspectrales: Indian-Pines et Salinas fournies par le capteur AVIRIS.

Le reste de ce chapitre est organisé comme suit. La section suivante présente les travaux liés aux approches Wrapper pour la sélection des attributs. Dans la section 3, nous présentons l'approche proposée en se basant sur la combinaison de l'information mutuelle normalisée NMI et la probabilité d'erreur PE avec le classifieur SVM. Les résultats obtenus et leur discussion sont présentés dans la section 4. Enfin, la section 5 conclut notre travail.

5.2 Les recherches antérieures

Rappelons que les méthodes de sélection d'attributs sont classées principalement en Filtre et Wrapper selon leur dépendance à l'algorithme de classification. Les approches Filtre sélectionnent les sous-ensembles d'attributs indépendamment de la performance de l'algorithme d'induction. Les approches Wrapper évaluent les sous-ensembles d'attributs durant la sélection en fonction de la précision de classification. De nombreux chercheurs ont démontré que les approches Wrapper peuvent donner de meilleures performances [Mao, 2004; Kohavi, 1997; Huang, 2006]. Parmi les méthodes de sélection d'attributs de type Wrapper présentées dans la littérature, Bermejo [Bermejo, 2014] propose l'optimisation (l'accélération) de la sélection des sous-ensembles en se basant sur l'approche Wrapper Incrémental avec le classifieur Naive Bayes, dans le cadre de la gestion efficace de la mémoire pour les très grands ensembles de données. Yu Enzhe [Yu, 2003] introduit une approche Wrapper basé sur les algorithmes génétiques GA-SVM pour la sélection de sous-ensembles d'attributs pour la vérification de l'identité selon la dynamique des frappes sur clavier ; il démontre la fiabilité de cette méthode et que le SVM représente un outil approprié aux approches Wrapper en termes de précision et de vitesse d'apprentissage. En télédétection, Li Zhuo [Li, 2008] propose une autre méthode Wrapper basée sur l'Algorithme Génétique GA et les machines à vecteurs de support SVM, pour la sélection d'attributs et la classification des données hyperspectrales. La méthode GA-SVM est appliquée à une image capturée par le

capteur HYPERION qui couvre une partie de la ville de Guangzhou, Chine ; il montre que cette méthode réduit significativement le coût de calcul tout en améliorant la précision de la classification.

Dans notre travail, nous proposons une nouvelle approche Wrapper basée sur l'utilisation de l'information mutuelle normalisée NMI avec la probabilité d'erreur PE. Les performances du classifieur SVM sont incorporées dans le processus de sélection des bandes pertinentes. Le but est l'amélioration des cartes thématiques classifiées.

5.3 Méthodologie proposée

5.3.1 Principe

Ce travail propose une nouvelle méthodologie Wrapper pour sélectionner les bandes les plus informatives parmi les images hyperspectrales utilisées. L'organigramme détaillé de la méthodologie proposée peut être schématisé comme suit :

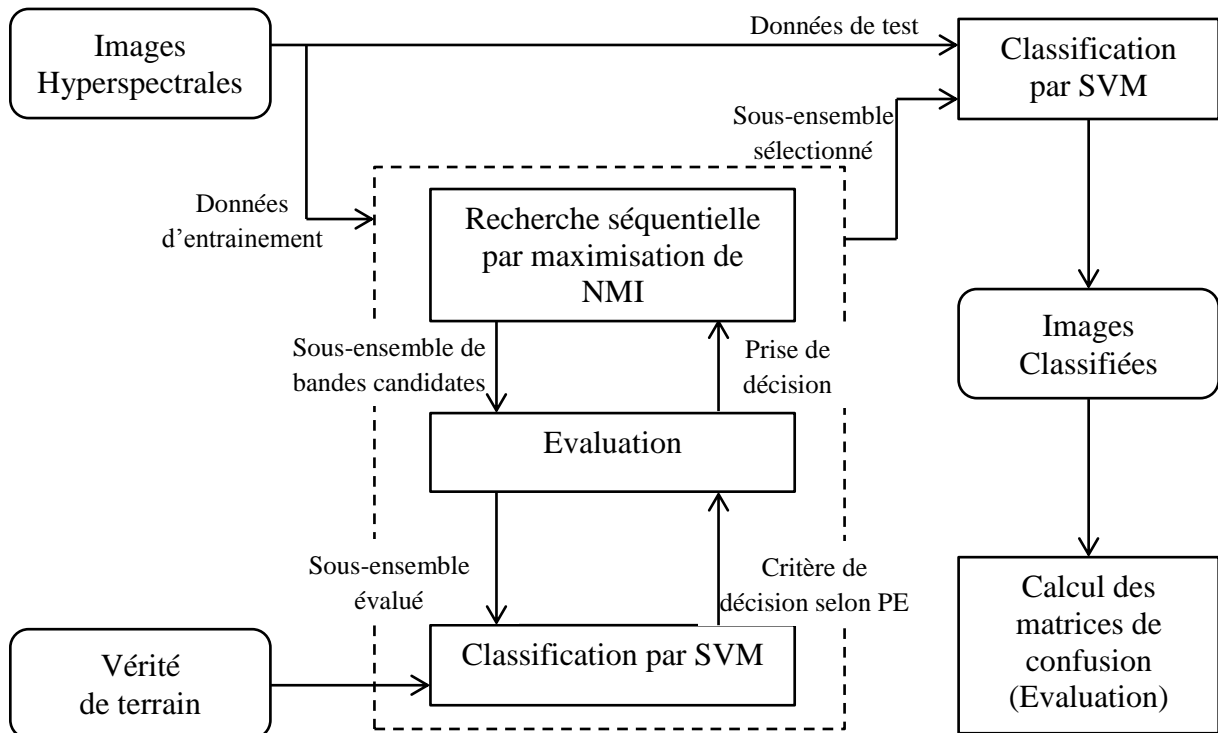


Figure 5. 1: Organigramme de la méthode Wrapper proposée.

L'algorithme proposé est basé essentiellement sur les trois étapes suivantes :

- Calculer l'information mutuelle normalisée NMI des bandes B avec la vérité de terrain GT en utilisant l'équation suivante, d'après l'équation 4.7 (chapitre4):

$$NMI(GT, B) = \frac{H(GT)+H(B)}{H(GT,B)} \quad (5.1)$$

- Utiliser l'approche Wrapper avec l'algorithme d'induction SVM-RBF pour construire le sous-ensemble réduit des bandes.
- Appliquer la probabilité d'erreur PE comme critère d'évaluation et de décision pour améliorer les performances de la classification

La nouveauté de notre contribution réside dans l'utilisation de l'information mutuelle normalisée NMI dans une approche Wrapper pour la réduction de la dimensionnalité et la classification des images hyperspectrales à l'aide des machines à vecteurs de support SVM. La seconde nouveauté est l'utilisation de la probabilité d'erreur avec la NMI comme critère d'évaluation et du contrôle de la redondance des bandes sélectionnées.

5.3.2 Probabilité d'erreur PE

Le succès d'un algorithme de sélection d'attributs pour les problèmes de classification peut être déterminé par la probabilité d'une mauvaise classification (misclassification) ; l'algorithme doit donner une erreur minimale de classification lorsque les attributs sont bien sélectionnés. Nous avons établi dans les chapitres précédents que les attributs ayant une valeur d'IM ou NMI élevée avec la vérité de terrain sont pertinents pour la classification [Guo, 2006 ; Sarhrouni, 2012 ; Nhaila, 2018b]. Dans ce contexte, l'erreur de Bayes PE justifie les avantages de l'utilisation de l'IM. L'erreur de Bayes a été majorée par le théorème de Hellman [Hellman, 1970] comme suit :

$$PE \leq \frac{H(X|Y)}{2} = \frac{H(X) - I(X,Y)}{2} \quad (5.2)$$

D'autre part, le théorème de Fano, connu aussi sous le lemme de Fano [Fano, 1961; Cover, 1991] donne la borne inférieure de PE :

$$\frac{H(X|Y) - 1}{\log_2(|X|)} = \frac{H(X) - I(X,Y) - 1}{\log_2(|X|)} \leq PE \quad (5.3)$$

En effet, le théorème de Fano relie la probabilité d'erreur à l'entropie conditionnelle : pour tout estimateur $\hat{X}: X \rightarrow \hat{X}$ avec $PE = \Pr\{X \neq \hat{X}\}$ nous avons :

$$H(PE) + PE \log|\chi| \geq H(X|\hat{X}) \geq H(X|Y) \quad (5.4)$$

Avec $|\chi|$ est le nombre de classe dans la vérité de terrain GT pour notre application.

Cette inégalité est majorée pour :

$$1 + PE \log|\chi| \geq H(X|Y) \quad (5.5)$$

D'où:
$$PE \geq \frac{H(X|Y) - 1}{\log |\chi|} \quad (5.6)$$

Puisque :
$$H(X|Y) = H(X) - I(X, Y) \quad (5.7)$$

Donc :
$$PE \geq \frac{H(X) - I(X, Y) - 1}{\log |X|} \quad (5.8)$$

L'expression finale de l'erreur de Bayes est :

$$\frac{H(X|Y) - 1}{\log |X|} = \frac{H(X) - I(X, Y) - 1}{\log |X|} \leq PE \leq \frac{H(X|Y)}{2} = \frac{H(X) - I(X, Y)}{2} \quad (5.9)$$

5.3.3 Application aux IHS

Dans notre approche Wrapper, nous exploitons l'intervalle de l'erreur de bayes, explicitée dans [Cover, 1991; Fano, 1961; Hellman, 1970 ; Kwak, 2007], pour déterminer l'efficacité d'une bande à la classification. Soit : $X=GT$ la vérité de terrain. $Y =GT_{est}$ est l'estimation de la vérité du terrain générée à partir du sous-ensemble des bandes à tester S .

$$\frac{H(GT|GT_{est}) - 1}{\log N_c} \leq PE \leq \frac{H(GT|GT_{est})}{2} \quad (5.10)$$

D'où:

$$\frac{H(GT) - I(GT, GT_{est}) - 1}{\log(N_c)} \leq PE \leq \frac{H(GT) - I(GT, GT_{est})}{2} \quad (5.11)$$

Dans notre processus de sélection, puisque N_c (Nombre de classe de la vérité de terrain) et $H(GT)$ sont constants, lorsque $IM(GT, GT_{est})$ est maximale, PE devient alors minimale. Ainsi, la bande candidate qui minimise la probabilité d'erreur présente une grande similarité avec la vérité-terrain. Pour le contrôle de la redondance, nous introduisons un seuil Th : Une bande sera retenue dans la sélection finale si elle minimise PE d'une valeur Th par rapport à sa dernière valeur. Le seuil Th est fixé à 0.001.

La Figure 5.2 donne la probabilité d'erreur PE calculée, d'après l'équation 5.11, pour chaque bande seule avec ($S=bande_{i=1 \text{ à } 220}$) :

$$\frac{H(GT) - IM(GT, GT_{est}) - 1}{\log(N_c)} \leq PE \leq \frac{H(GT) - IM(GT, GT_{est})}{2}$$

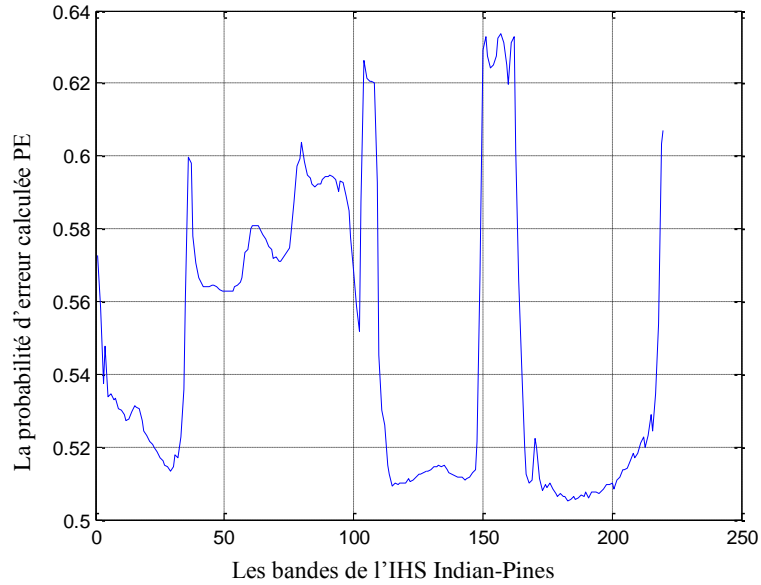


Figure 5. 2: La probabilité d'erreur calculée avec l'IM pour chaque bande avec GT de l'IHS Indian

Dans nos études, la comparaison de l'information mutuelle et l'information mutuelle normalisée nous a conduit à établir un nouvel intervalle PE* définit tel que :

$$\frac{H(GT) - NMI(GT, GT_{est}) - 1}{\log(Nc)} \leq PE^* \leq \frac{H(GT) - NMI(GT, GT_{est})}{2} \quad (5.12)$$

Avec la $NMI(GT, GT_{est})$ calculée par l'équation 5.1 :

$$NMI(GT, GT_{est}) = \frac{H(GT) + H(GT_{est})}{H(GT, GT_{est})} \quad (5.13)$$

La Figure 5.3 donne l'intervalle PE* calculé, d'après l'équation 5.12, pour chaque bande seule avec (S=bande $i=1$ à 220) :

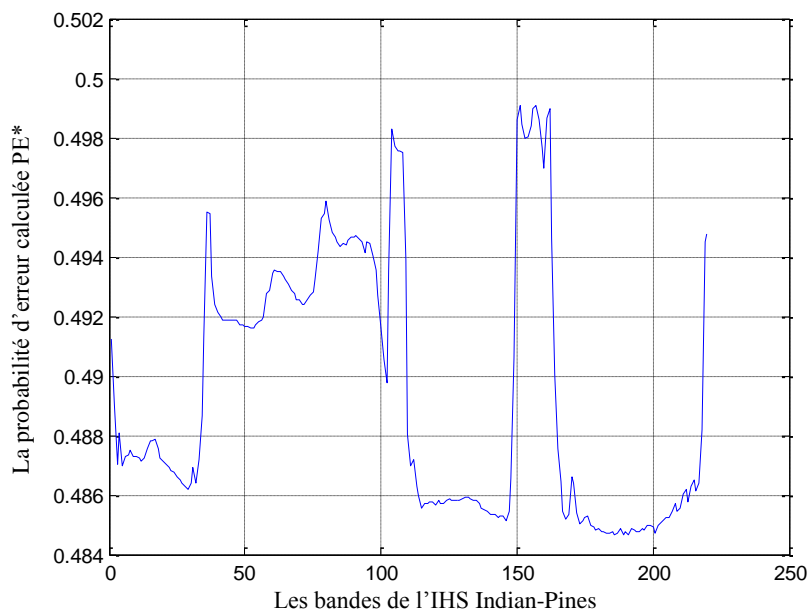


Figure 5. 3: L'intervalle PE* calculé avec la NMI pour chaque bande avec GT

Nous procédons à la comparaison des résultats de sélection des bandes pour les deux intervalles PE et PE* dans nos expérimentations.

5.3.4 L'algorithme proposé

L'idée principale de la méthode proposée est que la bande candidate B est une bonne approximation de la vérité-terrain GT si elle a une valeur plus élevée d'information mutuelle normalisée. Puisque notre algorithme est un schéma Wrapper, cette bande doit diminuer la PE calculée pour augmenter la précision de la classification. Notre algorithme est donc considéré comme une sélection de sous-ensembles de type Wrapper incrémental (IWSS) [Bermejo, 2009] utilisant l'information mutuelle normalisée [Nhaila, 2018c].

Notons que l'algorithme d'induction utilisé pour la classification est le SVM-RBF « les machines à vecteurs de support » avec noyau RBF [Nhaila, 2018a].

Le processus de sélection complet de la méthode proposée est le suivant:

Algorithme

Entrée: Image hyperspectrale

GT: Vérité de terrain;

B:={Bandes};

Sortie:

Le sous ensemble des bandes sélectionnées.

Carte Classifiée.

1 (Initialisation)

B ← les bandes de l'IHS

S ← {} // le groupe des bandes sélectionnées

L ← Nombre des bandes à sélectionner

Th ← Seuil de contrôle de redondance

2) Pour toutes B_i dans B, Calculer:

$$NMI(GT, B_i) = \frac{H(GT)+H(B_i)}{H(GT, B_i)} \quad // \text{ Avec } H(GT) \text{ est l'entropie de } GT$$

3) Processus de sélection

Sélectionner la bande B_{max} = argmax(NMI(B))

B ← B \ {B_{max}}; S ← {B_{max}}; G_{-est} = SVM(B_{max})

$$PE(S) = \frac{H(GT) - MI(GT, GT_{-est})}{2} - \frac{H(GT) - MI(GT, G_{-est}) - 1}{\log N_c} // \text{ ou } PE^*(S) = \frac{H(GT) - NMI(GT, GT_{-est})}{2} -$$

$$\frac{H(GT) - NMI(GT, G_{-est}) - 1}{\log N_c}$$

Tant que |S| < L Faire

B_{max} = argmax (NMI(B\S))

S ← S ∪ {B_{max}};

GT_{-est} = SVM(S)

PE⁺(S) // ou PE⁺*

```
Si  $PE^+ \leq PE - Th$  alors  
     $PE \leftarrow PE^+ //$  ou  $PE^* \leftarrow PE^{**}$   
Sinon  $S \leftarrow S \setminus \{B_{max}\};$   
       $B \leftarrow B \setminus \{B_{max}\};$   
Fin Si
```

Fin Tant que

4) S contient les bandes sélectionnées

5) Carte classifiée=SVM(S)

5.4 Expériences

5.4.1 Protocol expérimental

Dans ce chapitre, les deux images hyperspectrales d'AVIRIS Indian-Pines et Salinas, décrites au premier chapitre, sont utilisées dans les expériences. Elles ont des caractéristiques différentes en termes de dimensions et de type de classes.

Les performances de l'algorithme proposé sont évaluées en termes de réduction de dimensionnalité et de précision de la classification. Les métriques utilisées sont les suivantes: AA (la moyenne des classifications des classes), OA (le précision globale de classification), Kappa (le coefficient de concordance) et ICA (la classification individuelle de chaque classe). Les tests sont effectués avec un PC 64-b avec CPU Duo quad-core, 3 Go de RAM et 2.1Ghz.

La classification est réalisée avec le SVM-RBF [Nhaila, 2018a], les échantillons utilisés pour l'apprentissage et le test sont choisis d'une manière aléatoire. Trois cas ont été considérés dans notre étude: 10%, 25% et 50% des pixels de chaque donnée ont été utilisés pour l'entraînement et les échantillons restants, respectivement de 90%, 75% et 50%, ont été utilisés pour tester les modèles.

Afin de valider les résultats obtenus de notre méthode proposée (WNMIPE), nous les avons comparés à deux autres approches filtres qui sont l'algorithme de pertinence maximale-redondance minimale MRMR [Peng, 2005] et la sélection par maximisation de l'information mutuelle normalisée NMIBS [Nhaila, 2018b].

5.4.2 Résultats et discussion

Les résultats expérimentaux de l'approche proposée pour les bases de données Indian-Pines et Salinas sont présentés dans cette section et sont évalués à l'aide de quatre paramètres d'évaluation : ICA, AA, OA et le coefficient Kappa k.

✓ **Résultats de classification en fonction de la taille des échantillons d'entraînement.**

Cette étape consiste à sélectionner les bandes pertinentes parmi les 224 bandes des IHS d'entrées. Leurs classifications se font selon trois différentes tailles des échantillons d'entraînement (10%, 25% et 50%).

Le tableau 5.1 montre les résultats de la classification en termes des coefficients AA, OA et Kappa obtenus par la méthode proposée : un ensemble réduit de 49 bandes est sélectionné de l'IHS Indian-Pines.

Tableau 5. 1: AA(%), OA(%) et le coefficient kappa obtenus par la méthode proposée sur l'image Indian-Pines en utilisant 49 bandes sélectionnées avec trois cas d'ensembles d'apprentissage.

	Ensemble d'apprentissage 10%	Ensemble d'apprentissage 25%	Ensemble d'apprentissage 50%
AA	81.10	81.17	81.61
OA	87.61	88.58	90.29
Kappa	0.8679	0.8782	0.8964

Le tableau 5.2 montre les résultats de la classification en termes des coefficients AA, OA et Kappa obtenus par la méthode proposée : un ensemble réduit de 37 bandes est sélectionné de l'IHS Salinas.

Tableau 5. 2: AA(%), OA(%) et le coefficient kappa obtenus par la méthode proposée sur l'image Salinas en utilisant 37 bandes sélectionnées avec trois cas d'ensembles d'apprentissage.

	Ensemble d'apprentissage 10%	Ensemble d'apprentissage 25%	Ensemble d'apprentissage 50%
AA	96.46	97.44	97.90
OA	92.55	94.44	95.56
Kappa	0.9206	0.9407	0.9526

Les tableaux 5.1 et 5.2 nous permettent de formuler les observations suivantes:

- La méthode proposée fournit des résultats satisfaisants de classification avec une précision globale OA qui atteint 87,61 %, 88,58 % et 90,29 % pour respectivement 10 %, 25 % et 50 % comme échantillons d'apprentissage sur Indian-Pines. Dans le cas de l'IHS Salinas, on obtient 92,55%, 94,44% et 95,56%, avec respectivement 10%, 25% et 50% de données d'entraînement.

- Les résultats sont obtenus avec un nombre réduit de 49 bandes sélectionnées pour Indian-Pines et 37 bandes pour Salinas. Ceci nous permet de valider la sélection efficace de l'approche Wrapper proposée.
- Les résultats montrent également l'effet du nombre de pixels d'apprentissage utilisés pour la classification. Les paramètres (AA, OA et kappa) augmentent légèrement avec la taille des ensembles d'apprentissage. L'algorithme proposé fournit des résultats satisfaisants lorsqu'on réduit la taille des échantillons d'entraînement : c'est le cas de 10% où l'OA et le kappa atteignent respectivement 87,61% et 0,8679 pour Indian-Pines. Pour Salinas, on obtient 92,55% en OA et 0,9206 en coefficient kappa.

✓ Comparaison de l'algorithme proposé avec les méthodes MRMR et NMIBS.

Dans cette section, nous présentons les résultats de la méthode proposée en comparaison avec deux autres algorithmes Filtres de sélection de bandes: MRMR [Peng, 2005] et NMIBS [Nhaila, 2018b]. Les méthodes comparées utilisent les mêmes ensembles d'apprentissage et de test de 50% avec le classifieur SVM-RBF. Les résultats obtenus sont illustrés dans le tableau 5.3 pour la scène Indian-Pines et dans le tableau 5.4 pour Salinas.

La première colonne des tableaux 5.3 et 5.4 indique le numéro et le nombre total de pixels dans chaque classe des IHS. Les autres colonnes représentent les résultats obtenus des différentes méthodes. Les lignes représentent l'ICA de chaque classe des scènes. Les dernières lignes contiennent respectivement les coefficients AA, OA et Kappa.

D'après les deux tableaux 5.3 et 5.4, on peut faire les remarques suivantes :

- Pour Indian-Pines, le tableau 5.3 montre que la méthode proposée donne de meilleurs résultats que les autres méthodes avec 85,05 % de AA, 93,34 % d'OA et 0,9290 de Kappa. Le WNMIPÉ* proposé dépasse (p. ex. en termes de précision globale OA) le MRMR avec 9,39 %, le NMIBS avec 4.71% et le WNMIPÉ avec 4.17%.
- L'efficacité de la méthode proposée en utilisant l'IHS Indian-Pines est d'avantage illustrée à la figure 5.4 qui montre la vérité de terrain en (a) et les cartes classifiées obtenues en utilisant les algorithmes: MRMR en (b) avec OA de 83.95%, l'algorithme NMIBS en (c) avec OA égal à 88.63%, l'algorithme proposé WNMIPÉ en (d) avec OA égal à 89.17% et l'algorithme proposé WNMIPÉ* en (e) avec OA de 93.34%.

Tableau 5. 3: Les performances de la classification de l'IHS Indian-Pines en utilisant la méthode proposée en comparaison avec les algorithmes MRMR et NMIBS pour 60 bandes sélectionnées.

Les classes		Comparaison des Approches			
		MRMR	NMIBS	WNMIPE	WNMIPE*
1	54	82.61	78.26	78.26	82.61
2	1434	71.41	70.99	71.83	83.26
3	834	78.90	78.18	77.94	83.21
4	234	60.68	61.54	64.10	72.65
5	497	84.55	85.77	86.99	91.87
6	747	90.78	91.34	92.74	95.81
7	26	46.15	46.15	53.85	69.23
8	489	95.10	94.29	95.10	96.33
9	20	60.00	70.00	100	80.00
10	968	76.65	76.03	78.51	86.36
11	2468	81.04	80.71	81.77	87.76
12	614	81.11	77.20	82.08	83.39
13	212	96.12	98.06	98.06	98.06
14	1294	94.74	94.44	94.74	96.29
15	380	48.80	56.02	52.41	62.65
16	95	93.48	93.48	93.48	91.30
AA		70.11	78.28	81.37	85.05
OA		83.95	88.63	89.17	93.34
Kappa		0.8288	0.8787	0.8844	0.9290

Tableau 5. 4: Les performances de la classification de l'IHS Salinas en utilisant la méthode proposée en comparaison avec les algorithmes MRMR et NMIBS pour 37 bandes sélectionnées.

Les classes		Comparaison des Approches			
		MRMR	NMIBS	WNMIPE	WNMIPE*
1	2009	99.20	99.60	99.50	100
2	3726	99.95	99.89	100	100
3	1976	96.36	98.28	98.28	99.90
4	1394	99.26	99.71	99.85	99.56
5	2678	96.09	99.32	99.55	100
6	3959	99.65	99.95	99.95	99.85
7	3579	99.33	100	100	99.83
8	11271	89.98	86.43	86.43	91.76
9	6203	99.45	99.84	99.87	99.94
10	3278	86.27	96.52	97.25	98.72
11	1068	86.52	96.25	97.00	100
12	1927	99.79	99.79	99.90	100
13	916	97.82	99.34	99.34	99.13
14	1070	92.90	98.50	98.50	99.44
15	7268	44.44	59.41	60.31	78.27
16	1807	98.66	99.22	99.11	100
AA		92.85	95.75	95.93	97.90
OA		88.71	91.41	91.63	95.56
Kappa		0.8795	0.9083	0.9108	0.9526

- Nous constatons l'efficacité des approches Wrapper proposées WNMPIE et WNMPIE*. C'est une amélioration de l'approche Filtre utilisant l'Information Mutuelle Normalisée [Nhaila, 2018b]: l'inclusion des performances du classifieur SVM-RBF dans le processus de sélection à travers le contrôle de la probabilité d'erreur PE augmente la pertinence des bandes sélectionnées. La visualisation des cartes classifiées en figure 5.4 (d et e) confirme l'avantage d'utiliser l'approche Wrapper avec sélection à base de NMI et le contrôle de PE.

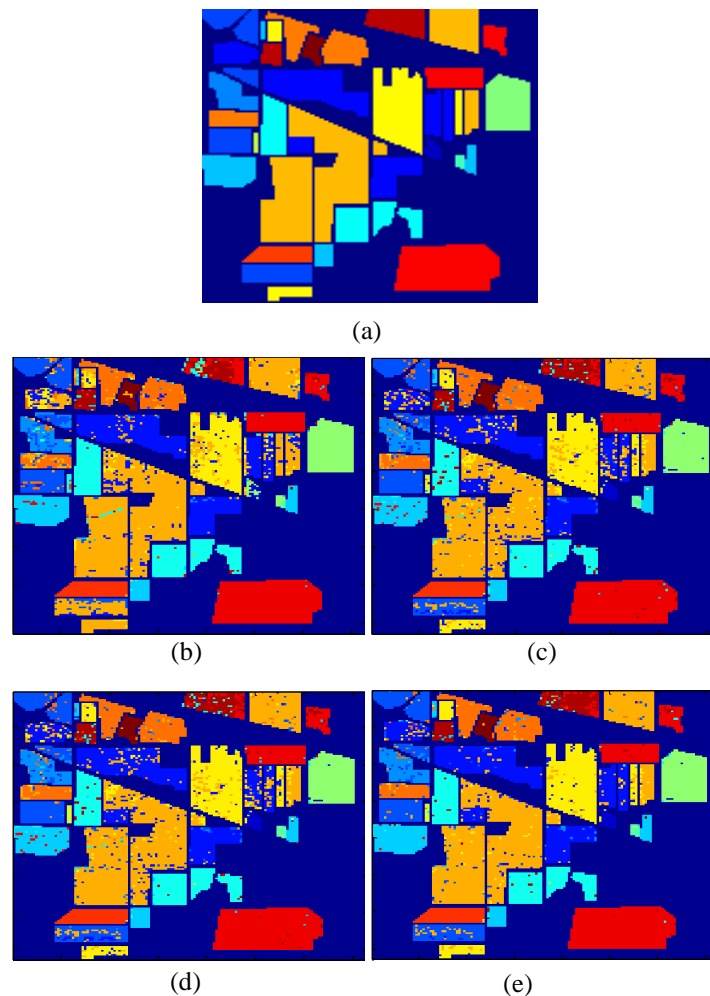


Figure 5. 4 : *L'image hyperspectrale Indian-Pines: (a) Carte de la vérité de terrain ; et les cartes classifiées obtenues pour 60 bandes sélectionnées en utilisant les algorithmes (b) MRMR (OA=83.95%), (c) NMIBS (88.63%) et les méthodes proposées WNMPIE en (d) avec (OA=89.17%) et WNMPIE* (e) avec (OA=93.34%).*

- Pour la scène de Salinas et à partir du Tableau 5.4, nous pouvons remarquer que la méthode proposée WNMPIE/WNMPIE* surpasse les autres méthodes en termes de AA, OA et Kappa. L'ICA de 12 classes (1, 2, 3, 5, 8, 9, 10, 11, 12, 14, 15 et 16), parmi

les 16 classes de l'IHS, augmente significativement. Dans le cas échéant, l'ICA des quatre classes restantes (4, 6, 7 et 13) ne dépasse pas celui de l'approche proposée de plus de 0.21%. Dans l'exemple de la classe 11 nommée *laitue_romaine_4wk*, le taux de classification individuelle ICA augmente de (13,48%) entre MRMR (86,52%) et WNMIPÉ* (à 100%). L'exemple du cas échéant est l'ICA de la classe 13 nommée (*laitue_romaine_6wk*) qui diminue de 0,21% entre WNMIPÉ et WNMIPÉ*.

- Les cartes classifiées de Salinas en utilisant la méthode proposée WNMIPÉ* (e)/ WNMIPÉ (d), l'algorithme NMIBS (c) et MRMR (b) sont présentées à la figure 5.5. La comparaison avec la vérité de terrain, figure 5.5 (a), illustre la qualité des cartes reproduites et classifiées en utilisant l'approche proposée WNMIPÉ/WNMIPÉ*: un ensemble réduit de 37 bandes est suffisant pour distinguer la plupart des classes de la scène.

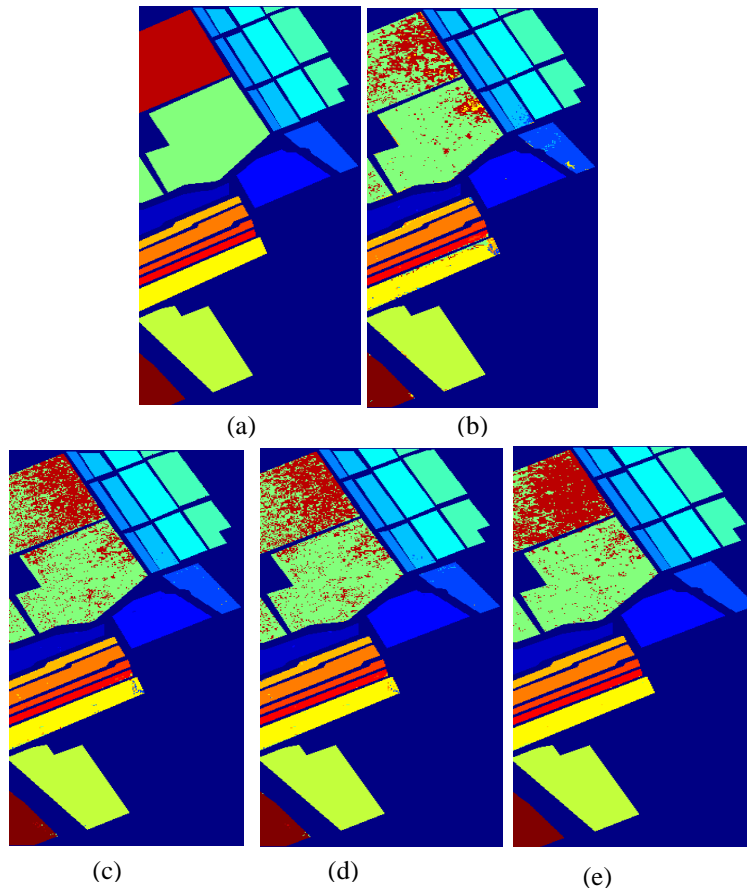


Figure 5. 5 : *L'image hyperspectrale Salinas: (a) Carte de la vérité de terrain ; et les cartes classifiées obtenues pour 37 bandes sélectionnées avec une précision globale OA (entre parenthèses) en utilisant les algorithmes reproduits (b) MRMR (88.71%), (c) NMIBS (91.41%), (d) WNMIPÉ (91.63%) et (e) WNMIPÉ* (95.56).*

5.5 Conclusion

L'objectif principal de ce chapitre est d'introduire une nouvelle approche Wrapper pour la réduction de dimensionnalité et la classification des images hyperspectrales. L'algorithme est basé sur la sélection des attributs en utilisant l'information mutuelle normalisée NMI, le contrôle de la probabilité d'erreur PE et le classifieur SVM-RBF. Dans le processus de sélection, chaque bande candidate devait vérifier deux critères : augmenter l'information mutuelle normalisée et minimiser la probabilité d'erreur pour être ajoutée à l'ensemble sélectionné.

Les expériences sont effectuées sur deux images hyperspectrales d'AVIRIS de la NASA: Indian-Pines et Salinas. Les résultats obtenus pour trois ensembles d'échantillons d'entraînement (10 %, 25 % et 50 %) ont été évalués à l'aide de plusieurs paramètres de classification (ICA, AA, OA et coefficient Kappa). L'analyse des résultats présentés et la visualisation des cartes classifiées montrent que les algorithmes proposés WNMPIE et WNMPIE* améliorent considérablement la précision de la classification par rapport aux autres méthodes MRMR et NMIBS pour les deux IHS.

Nous concluons que l'approche Wrapper proposée qui utilise conjointement l'information mutuelle normalisée pour la sélection des bandes pertinentes et la probabilité d'erreur PE pour le contrôle de la redondance donne de meilleures performances que les méthodes utilisant l'information mutuelle normalisée dans des approches Filtres. Cependant les approches filtres demandent moins de temps d'exécution que les approches Wrapper qui font appel à l'algorithme d'induction dans le processus de sélection.

Dans le chapitre suivant, nous utilisons l'approche Filtre pour montrer l'effet de l'ajout de l'information spatiale à l'information spectrale pour la réduction de dimensionnalité des IHS.



Chapitre
6

Apport de la texture pour la réduction de dimensionnalité et la classification des images hyperspectrales : une nouvelle méthode basée sur les caractéristiques GLCM et l'information mutuelle

Résumé :

La réduction de dimensionnalité se fait par sélection ou extraction d'attributs. Contrairement à la sélection, l'extraction d'attributs modifie l'ensemble de données d'origine pour définir un nouveau sous-ensemble de caractéristiques. Dans ce chapitre, nous proposons une nouvelle méthode de réduction de dimensionnalité et de la classification des IHS qui combine l'extraction de l'information spatiale de texture et l'information spectrale pour la sélection de bandes. La classification se fait sur les bandes originales sélectionnées en utilisant les machines à vecteurs de support SVM. L'information spatiale de texture est extraite de la matrice de cooccurrence des niveaux de gris (GLCM) : nous utilisons l'Homogénéité, le Contraste, la Corrélacion et l'énergie. Les expériences sont effectuées sur les trois images hyperspectrales déjà présentées dans les chapitres précédents : Indian-Pines, Salinas et Pavia-University. Les résultats obtenus montrent que de la fusion de l'information spectrale et spatiale augmente la précision de la classification dans un bon intervalle de temps.

6.1 Introduction

Les images hyperspectrales consistent à acquérir des spectres pour tous les pixels de l'image, elles fournissent plus d'une centaine de bandes d'une même région avec des informations plus détaillées, la riche disponibilité des données hyperspectrales augmente la discrimination des signatures spectrales par rapport aux images multispectrales. Pour cet effet, plusieurs travaux ont été développés pour éliminer les bandes redondantes ou non pertinentes par sélection ou extraction de bandes, notamment dans les applications d'analyse de la couverture terrestre; Sarhrouni [Sarhrouni, 2012] a proposé une stratégie Filtre basée sur l'information mutuelle pour la sélection de bandes nommée (MIBS). Un nouvel algorithme basé sur l'optimisation pour sélectionner les bandes spectrales des IHS a été proposé par Medjahed [Medjahed, 2016] en utilisant « Gray Wolf Optimizer » (GWO). Ces deux méthodes sont basées uniquement sur des informations spectrales. D'autre part, Zhu [Zhu, 2016] introduit une sélection non supervisée des bandes par extraction des ensembles dominants (DSEBS). Elle consiste à analyser la consistance spatiale-Spectrale pour définir les bandes informatives. Notre contribution consiste à proposer un nouvel algorithme, nommé (GLMI), qui combine les caractéristiques extraites de la matrice de cooccurrence des niveaux de gris GLCM avec l'information mutuelle IM. L'objectif est la sélection des bandes pertinentes nécessaires à la classification des images hyperspectrales [Nhaila, 2018d].

L'efficacité de l'algorithme proposé est évaluée en termes de différentes métriques d'évaluation sur les trois scènes hyperspectrales capturées par les capteurs ROSIS et AVIRIS. Le reste de ce chapitre est organisé comme suit. La deuxième section résume les méthodes de classification basées sur les textures et des travaux connexes. Dans la section 3, nous décrivons la méthodologie proposée en utilisant à la fois l'information mutuelle et les caractéristiques de texture. Les résultats expérimentaux sont présentés et discutés dans la section 4. Enfin, la section 5 conclut notre travail.

6.2 Les méthodes de classification basées sur la texture

La mesure de la texture est importante pour extraire l'information spatiale dans la classification des images. De nombreuses approches ont été appliquées pour l'analyse de texture et peuvent être classées en trois catégories : les méthodes spectrales, structurelles ou statistiques [Gonzalez, 2002], figure 6.1.

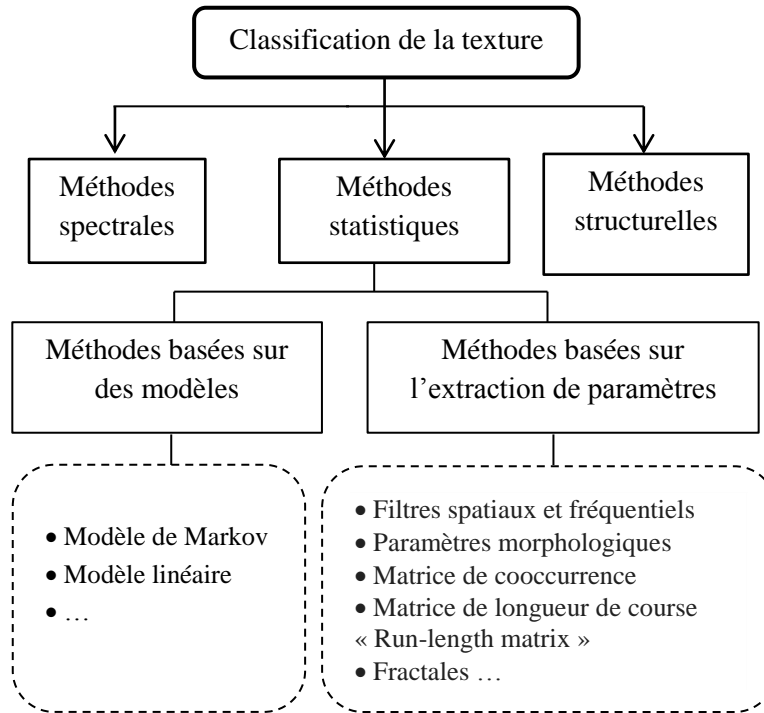


Figure 6. 1: *Méthodes de classification de texture*

Les méthodes spectrales consistent à convertir une image du domaine spatial en domaine fréquentiel et vice-versa en utilisant des filtres. De nombreux travaux ont été développés avec succès dans le contexte du traitement des images hyperspectrales [Rajadell, 2013 ; Chen, 2014]. Les études sur les méthodes structurelles sont plus complexes que les méthodes statistiques lorsque les textures sont faiblement structurées [Qian, 2013]. Les méthodes statistiques analysent la distribution spatiale des niveaux de gris sur la base des propriétés statistiques des images. On distingue deux familles des méthodes statistiques: la première regroupe les méthodes basées sur des modèles tels que le modèle de Markov [Li, 2014]. La deuxième famille des méthodes statistiques contient les méthodes basées sur l'extraction de paramètres telles que Fractals [Zhao, 2016] et les méthodes basées sur les algorithmes morphologiques [Song, 2014].

Dans cette étude, nous utilisons la matrice de cooccurrence pour extraire les caractéristiques de texture. Notre principale motivation derrière ce choix est due au fait qu'il s'agit de l'une des méthodes statistiques les plus utiles pour l'analyse de la texture des images depuis sa proposition par Haralick [Haralick, 1973; Agarwal, 2015]. La matrice de cooccurrence des niveaux de gris (GLCM) a été utilisée avec succès pour classifier la texture dans de nombreuses applications des IHS [Naganathan, 2008 ; Tsai, 2013].

6.3 Méthodologie proposée

6.3.1 Principe

Dans ce travail, nous introduisons l'apport de l'information spatiale des bandes dans le processus de sélection par information mutuelle pour la réduction de dimensionnalité et la classification des images hyperspectrales. Dans l'algorithme proposé (GLMI), l'information spatiale est incluse à travers quatre éléments de texture extraits de la matrice GLCM : nous utilisons l'énergie, l'homogénéité, le contraste et la corrélation. L'information mutuelle est ensuite appliquée pour sélectionner l'ensemble optimal de bandes spectrales à partir des images hyperspectrales d'entrée et éliminer celles qui sont redondantes en fonction des mesures de texture.

Le classifieur utilisé est les Machines à Vecteurs de Support SVM-RBF étudié au chapitre 3 [Nhaila, 2018a]. Pour valider l'efficacité de la méthode GLMI, nous utilisons les trois IHS Indian-Pines, Salinas et Pavia-University des capteurs AVIRIS et ROSIS.

L'organigramme de la méthodologie proposée est illustré dans la figure 6.2. Le processus détaillé de cette méthode est décrit dans cette section.

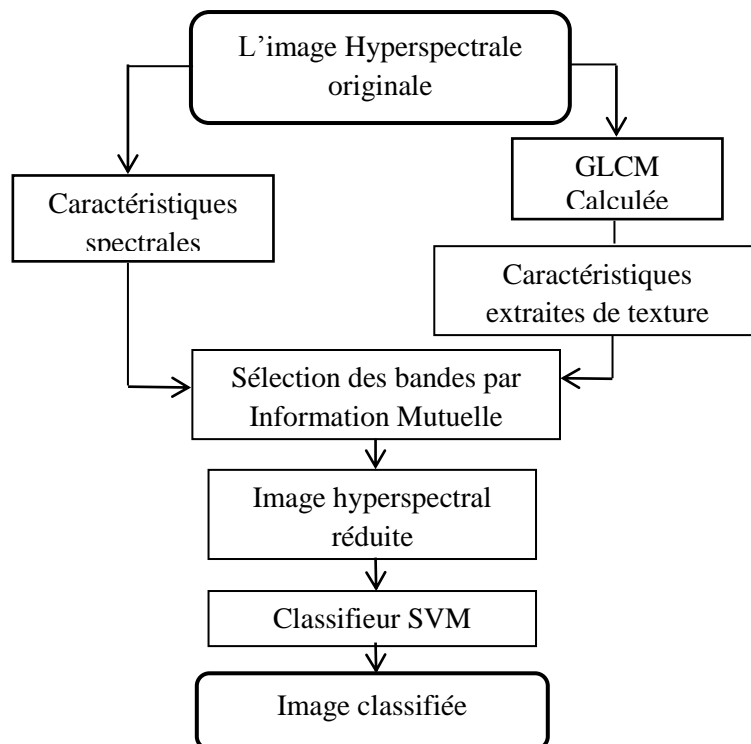


Figure 6. 2: L'organigramme de la méthodologie proposée

6.3.2 Sélection des bandes par IM et caractéristiques de texture

6.3.2.1 Rappel sur l'information mutuelle

L'information mutuelle IM est une mesure statistique de la similarité entre deux variables aléatoires : une référence (dans notre cas, la carte de vérité de terrain) que nous notons A et chaque bande notée B.

La IM entre A et B est donnée comme suit:

$$IM(A, B) = \sum \log_2 p(A, B) \frac{p(A, B)}{p(A)p(B)} \quad (6.1)$$

En relation avec l'entropie de Shanon, la IM peut être exprimée sous la forme :

$$IM(A, B) = H(A) + H(B) - H(A, B) \quad (6.2)$$

Parmi les méthodes de sélection des bandes par information mutuelle, Sarhrouni [Sarhrouni, 2012] présente un algorithme MIBS que nous reproduisons pour la comparaison de la sélection dans ce chapitre. Le processus de sélection de cet algorithme est décrit dans le chapitre 3.

6.3.2.2 Extraction des caractéristiques de texture

Les méthodes de classification des images hyperspectrales utilisent l'information spectrale pour la sélection des bandes. Les sous-ensembles produits sont utilisés comme entrée du classifieur. L'algorithme GLMI proposé exploite l'information spectrale des bandes en ajoutant la relation spatiale des pixels en utilisant la matrice de cooccurrence GLCM [Nhaila, 2018d].

La taille de la matrice GLCM est égale au nombre de niveaux de gris de l'image ; la distribution dépend de la distance d entre deux pixels dans quatre directions $\theta : 0^\circ, 45^\circ, 90^\circ$ et 135° . La figure 6.3, montre un exemple de calcul de la matrice de cooccurrence à partir d'une image 5 x 5 composée de 3 niveaux de gris (0,1,2) dans le cas de $d=(0,1)$.

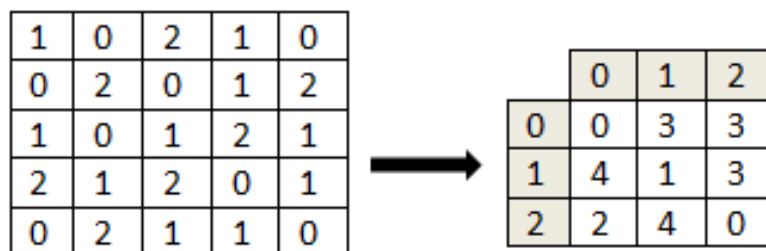


Figure 6. 3: Image 5x5 avec 3 niveaux de gris et la matrice de cooccurrence correspondante

Pour chaque bande de l'image hyperspectrale, nous calculons la matrice GLCM correspondante. Diverses caractéristiques peuvent être extraites de cette matrice. Dans notre cas, nous avons utilisé les caractéristiques suivantes : contraste, corrélation, énergie et homogénéité.

6.3.2.2.1 *Contraste*

Il mesure le contraste d'intensité entre deux pixels. Pour une image constante, le contraste est 0. Il est calculé à l'aide de l'équation 6.3.

$$\text{Contraste} = \sum_i \sum_j (i - j)^2 P(i, j) \quad (6.3)$$

Avec: $P(i, j)$: les éléments i, j de la GLCM

Illustration

La figure 6.4 représente le contraste et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes de contraste et d'IM. Les valeurs de la longueur d'onde centrale des bandes de l'IHS sont disponibles dans (https://engineering.purdue.edu/~biehl/MultiSpec/aviris_documentation.html).

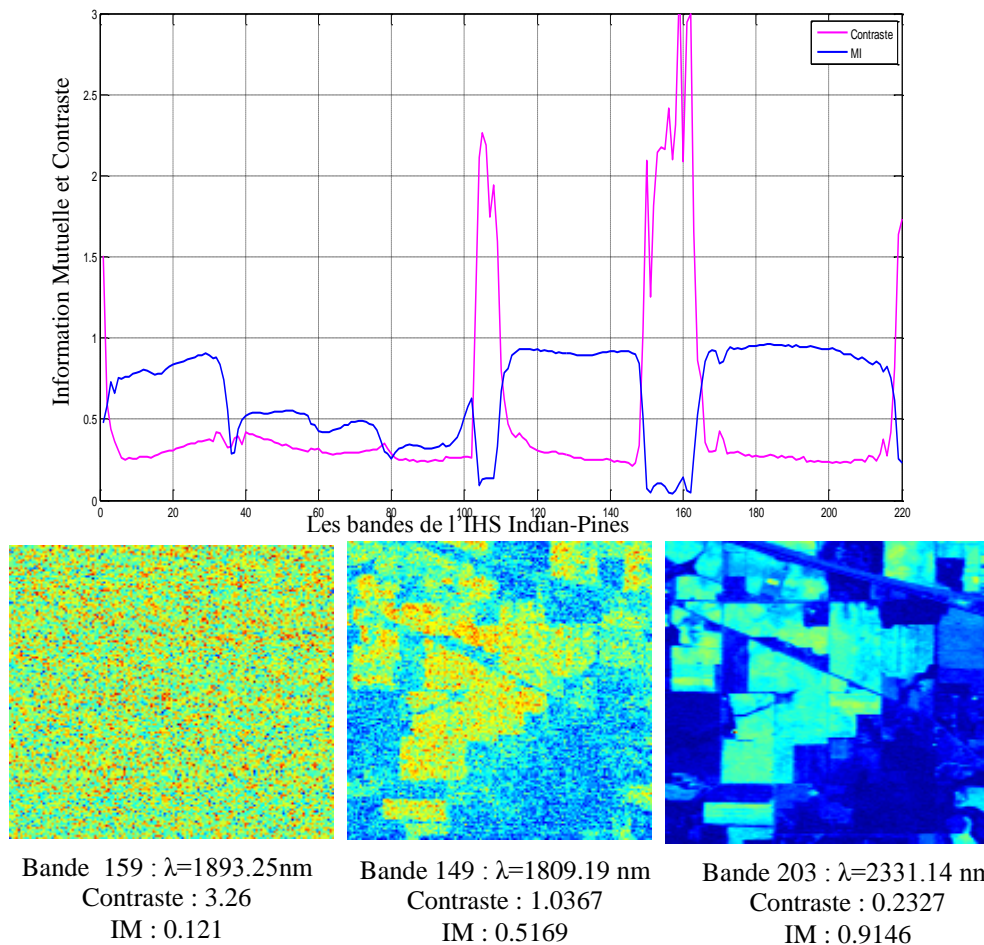


Figure 6. 4: Le contraste et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, avec trois bandes et leurs valeurs correspondantes de Contraste et d'IM

L'analyse de la figure 6.4 montre que :

- Lorsque le contraste est maximal ou trop élevé, nous avons des bandes trop bruitées et non porteuses d'information, c'est le cas de la bande 159.
- Les bandes qui ont un contraste moyen et dont l'IM est moyenne sont tel que l'information contenue est bruitée comme l'illustre la bande 149.
- Lorsque le contraste n'est pas trop faible et l'information mutuelle est importante, cela signifie que la bande est trop informative et contient moins de bruit. Voir la bande 203.
- Il n'y a pas de bandes qui ont une information mutuelle importante et en même temps un contraste élevé.
- Théoriquement, un contraste nul (trop faible) sera accompagné d'une information mutuelle trop faible : les classes sont inséparables.

Constat 1: Les bandes dont la valeur du contraste est moyenne et dont l'information mutuelle est importante sont les bandes les plus pertinentes.

6.3.2.2 Corrélation

La corrélation mesure la dépendance linéaire des niveaux de gris entre les pixels. C'est NaN pour une image constante.

$$Corrélation = \sum_i \sum_j \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} P(i, j) \quad (6.4)$$

Avec:

$P(i, j)$: les éléments i, j de la GLCM

μ : la moyenne de la GLCM

σ : L'écart-type.

Illustration

La figure 6.5 représente la corrélation et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes de corrélation et d'IM.

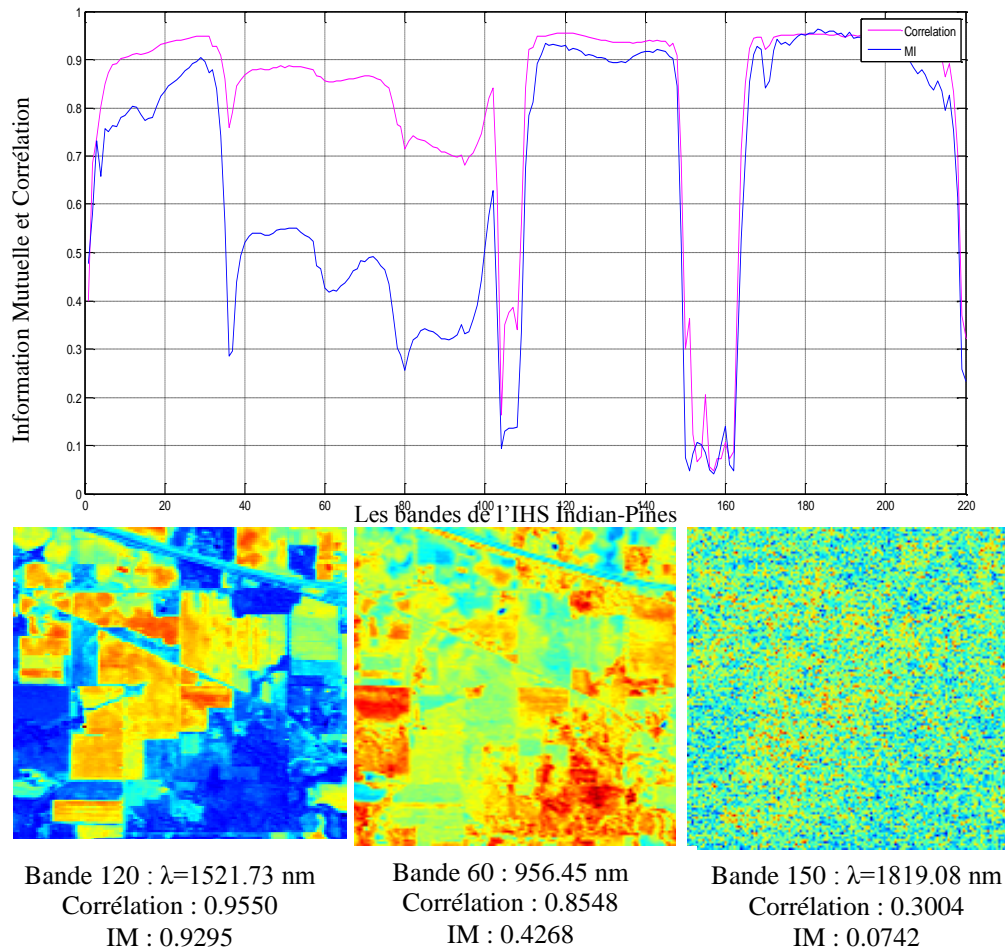


Figure 6.5 : La corrélation et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, avec trois bandes et leurs valeurs correspondantes de corrélation et d'IM

D'après l'analyse de la figure 6.5, on remarque que:

- Lorsque la corrélation est trop élevée, et l'information mutuelle est élevée, la bande est trop informative. La bande 120 en est un exemple. .
- Les bandes qui ont une corrélation élevée et dont l'IM est moyenne sont moins informatives que les bandes dont l'IM et la corrélation sont importantes. Voir exemple de la bande 60. La bande qui a une valeur maximale de corrélation signifie qu'il y a un regroupement des pixels qui se ressemblent, mais ne signifie pas la pertinence à la classification.
- Lorsque la corrélation et l'information mutuelle de la bande sont faibles cela signifie que la bande est n'est pas porteuse d'information. Voir la bande 150.

Constat 2: La corrélation maximale ne permet pas d'extraire les bandes les plus pertinentes sauf si ces bandes sont porteuses d'information (l'IM est élevée). L'IM et la corrélation représentent deux notions différentes : l'IM mesure la similarité de la bande avec la vérité de

terrain, par contre la corrélation mesure une caractéristique spatiale de la bande (la dépendance linéaire des niveaux de gris entre les pixels).

6.3.2.2.3 Énergie

L'énergie E mesure la somme des éléments carrés dans la GLCM en utilisant l'équation 6.5. E=1 si l'image est constante.

$$Energie = \sum_i \sum_j P(i, j)^2 \quad (6.5)$$

Avec:

$P(i, j)$: les éléments i, j de la GLCM

La figure 6.6 représente l'énergie et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'énergie et d'IM.

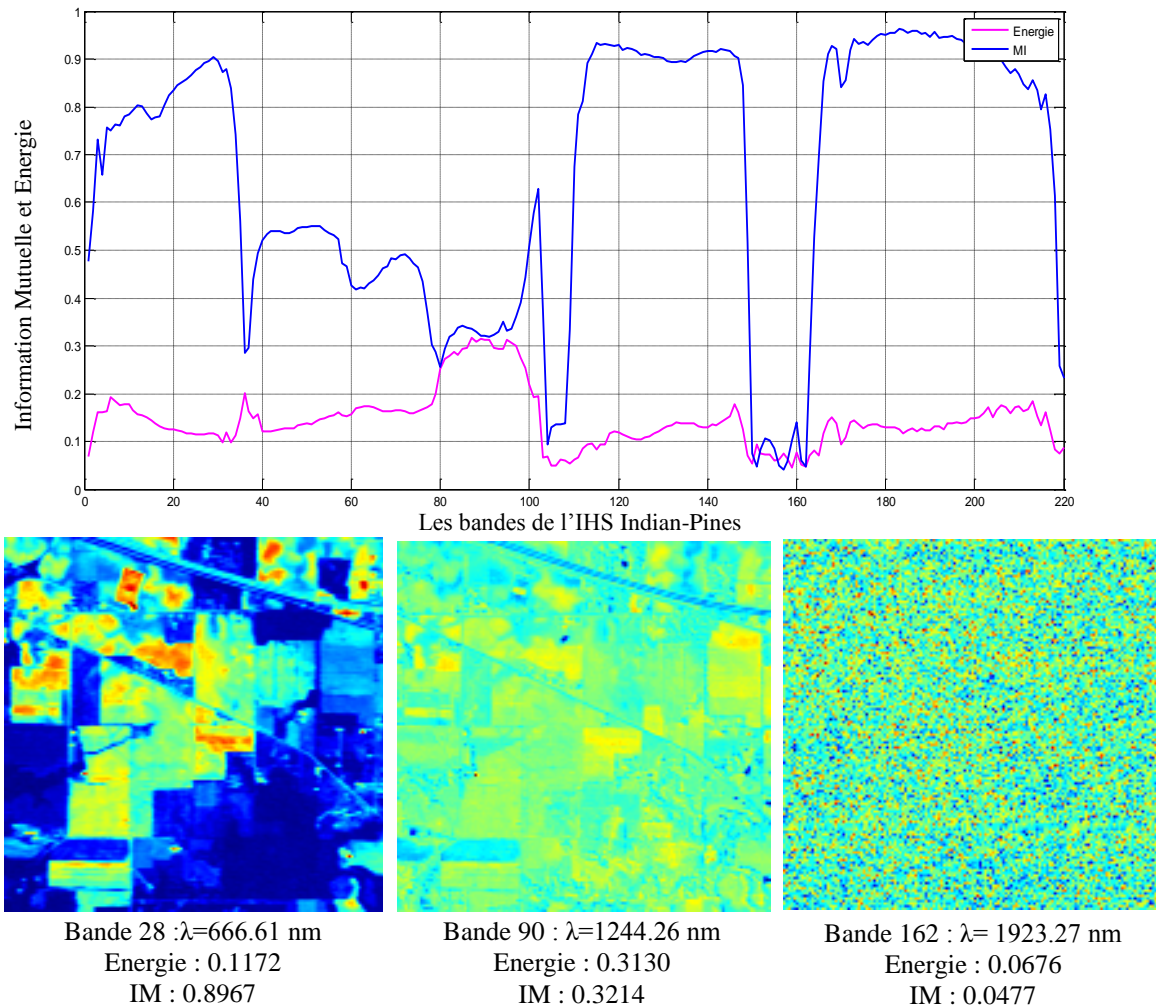


Figure 6. 6: L'énergie et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'énergie et d'IM.

L'analyse de la figure 6.6 montre que :

- Lorsque l'énergie est élevée, et l'information mutuelle est moyenne, la bande est moyennement informative. La bande 90 en est un exemple. .
- Les bandes qui ont une énergie moyenne et dont l'IM est élevé sont pertinentes pour la classification. Voir exemple de la bande 28.
- Lorsque l'énergie et l'information mutuelle de la bande sont faibles cela signifie que la bande n'est pas porteuse d'information. Voir la bande 162.

Constat 3: L'énergie toute seule ne permet pas de prélever les bandes les plus pertinentes. Les deux courbes de l'IM et l'énergie représentent deux notions différentes : l'IM mesure la similarité de la bande avec la vérité de terrain, par contre l'énergie mesure une propriété spatiale de la bande (c'est une relation entre les pixels). Une bande est plus pertinente pour la classification si elle est énergétique et ressemble à la vérité de terrain.

6.3.2.2.4 Homogénéité

L'homogénéité mesure la proximité de la distribution des éléments dans la diagonale du GLCM à l'aide de l'équation 6.6 suivante.

Elle a une valeur maximale lorsque tous les éléments de l'image sont identiques.

$$Homogénéité = \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i, j) \quad (6.6)$$

Avec:

$P(i, j)$: les éléments i, j de la GLCM

μ : la moyenne de la GLCM

σ : L'écart-type.

La figure 6.7 représente l'homogénéité et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'homogénéité et d'IM.

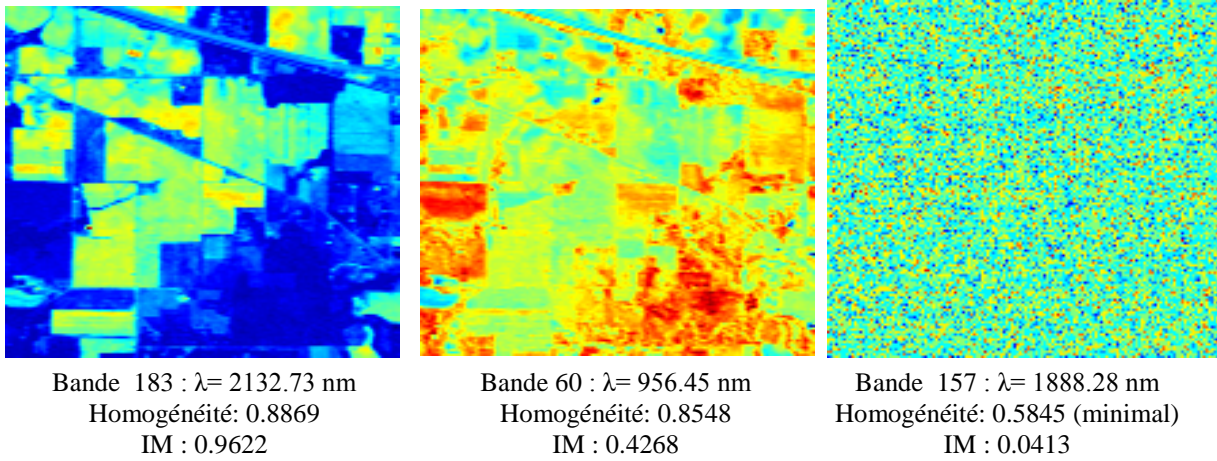
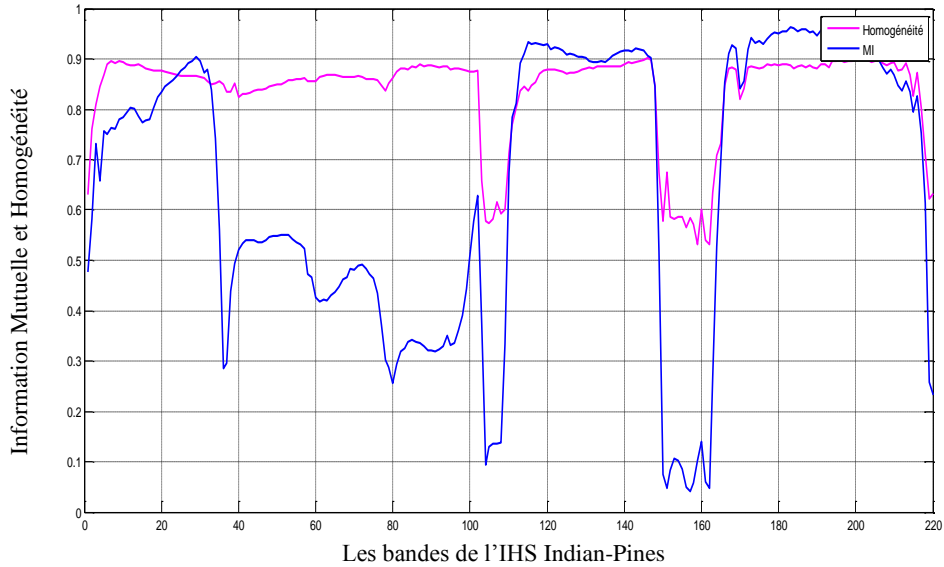


Figure 6. 7: L'homogénéité et l'information mutuelle des 220 bandes de l'IHS Indian-Pines, ainsi que trois bandes et leurs valeurs correspondantes d'homogénéité et d'IM.

D'après l'analyse de la figure 6.7, on remarque que :

- Lorsque l'homogénéité est élevée, et l'information mutuelle est moyenne, la bande est moyennement informative. La bande 60 en est un exemple. .
- Les bandes qui ont une l'homogénéité élevée et dont l'IM est élevé sont pertinentes pour la classification. Voir exemple de la bande 183.
- Lorsque l'homogénéité et l'information mutuelle de la bande sont faibles cela signifie que la bande n'est pas porteuse d'information. Voir la bande 157.

Constat 4: Une bande plus homogène contient trop de successions des pixels de même niveau de gris : il y a possibilité d'existence des classes (zone homogène). Cependant, l'homogénéité élevée d'une bande ne signifie pas qu'elle contient de l'information trop pertinente à la

classification. La combinaison de l'homogénéité et l'IM est importante pour sélectionner les bandes les plus pertinentes.

6.3.3 L'algorithme proposé

L'algorithme reproduit MIBS permettait de sélectionner les bandes pertinentes en fonction de l'information spectrale; mais il ne tient pas compte de l'information spatiale. Pour surmonter cet inconvénient, nous proposons un nouvel algorithme spectral-spatial nommé (GLMI). Dans cette méthode, nous utiliserons quatre caractéristiques spatiales qui caractérisent la texture extraite via la matrice de cooccurrence des niveaux de gris (GLCM). Nous utilisons les quatre caractéristiques suivantes: l'énergie, le contraste, l'homogénéité et la corrélation. Les expériences sont validées par le classifieur SVM-RBF. L'algorithme proposé est décrit comme suit:

L'Algorithme GLMI: Soit GLCM la matrice qui contient les quatre caractéristiques de texture des HIS: H est l'homogénéité, C le contraste, Cor la corrélation et E l'énergie. SS est l'ensemble des bandes déjà sélectionnées et S la bande candidate. SS est initialement vide. R est l'ensemble des bandes candidates, il contient initialement toutes les bandes. X est le nombre des bandes à sélectionner et TH c'est le seuil de contrôle de la redondance. Soit Critère =C ou Cor ou E ou H.

1) Features Extraction

Calculer la matrice GLCM pour toutes les bandes.

$C \leftarrow GLCM(1); Cor \leftarrow GLCM(2); E \leftarrow GLCM(3); H \leftarrow GLCM(4);$

2) Pour Critère $\in \{C, Cor, E, H\}$ Faire les étapes 3, 4 et 5.

3) Ordonner R dans l'ordre décroissant du Critère.

4) Sélection de la première bande

- Sélectionner la bande d'indice s tel que : $S = \max(R)$ et $s = \text{indice}(S)$
- $SS \leftarrow S; R \leftarrow R \setminus S; C_{est0} = Band(s);$
- $IM^* = \text{Information mutuelle}(Gt, C_{est0})$

5) Processus de sélection des bandes suivantes:

Tant que $|SS| < X$ faire

Sélectionner la bande $\text{index}_s S = C_{\max}(s) \text{ or } Cor_{\max}(s) \text{ or } E_{\max}(s) \text{ or } H_{\max}(s);$

$R \leftarrow R \setminus S;$

$C_{est} = \frac{C_{est0} + Band(S)}{2};$

$IM = \text{Information mutuelle}(Gt, C_{est})$

Si $IM > IM^* + Threshold$ **alors**

$$IM^* = IM; C_{-est0} = C_{-est}; SS \leftarrow SS \cup S;$$

Fin Si

Fin Tant que

6) L'ensemble SS contient les bandes sélectionnées selon le Critère (C, Cor, E et E) et selon le seuil de contrôle de redondance TH

6.4 Résultats expérimentaux

6.4.1 Protocol expérimental

Pour évaluer l'efficacité de l'algorithme proposé, l'application est faite sur les trois images hyperspectrales : Indian-Pines, Salinas et Pavia-University. Rappelons que les caractéristiques de ces images diffèrent selon la taille, le nombre de bandes et des classes ainsi que le type des zones des images.

La performance de la méthode proposée est évaluée en termes de réduction de la dimensionnalité et de précision de classification à l'aide de plusieurs paramètres d'évaluation: taux de classification Tx, classification individuelle de chaque classe (ICA), précision globale de classification (OA) et le coefficient kappa (k). Le temps d'exécution est également utilisé.

L'implémentation des programmes est faite sur un PC 64-b quad-core Duo PC CPU 2.1Ghz fréquence avec 3Go de RAM.

Pour les trois images hyperspectrales utilisées dans ces expériences, le classifieur utilisé est les Machines à Vecteur de Support SVM-RBF. Pour construire les modèles de classification et la comparaison avec d'autres algorithmes, 50 % et 10% des pixels étiquetés sont choisis au hasard pour être utilisés pour l'entraînement et le reste sera utilisé pour le test de classification.

6.4.2 Résultats et discussion

Les résultats expérimentaux des algorithmes MIBS et GLMI sur les trois IHS sont présentés et discutés en détail dans cette section.

6.4.2.1 Résultats et analyse de comparaison des algorithmes MIBS et GLMI selon le nombre de bandes sélectionnées et différents seuils de redondance

6.4.2.1.1 Tableaux et figures des résultats

Les tableaux 6.1, 6.2 et 6.3 donnent respectivement le taux de classification T_x pour les trois images Indian-Pines, Salinas et Pavia_University. Ils donnent la classification en fonction du nombre de bandes sélectionnées pour différents seuils de contrôle de redondance T_h . L'information spectrale (mesurée par la méthode MIBS) est comparée avec la combinaison de l'information spectrale et les caractéristiques spatiales (traitée par la méthode proposée GLMI).

Les graphiques des figures 6.8, 6.9 et 6.10 illustrent les résultats de classification obtenus pour différents nombres de bandes X retenues dans le cas de seuil de redondance $T_h = -0.02$ pour respectivement les trois IHS Indian-Pines, Salinas et Pavia-University.

Tableau 6. 1: Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Indian-Pines.

Nombre de bandes		Le taux de classification (%) selon les différents seuils																			
		-0.02					-0.01					-0.005					0				
		MIBS	GLMI				MIBS	GLMI				MIBS	GLMI				MIBS	GLMI			
			C	Cor	E	H		C	Cor	E	H		C	Cor	E	H		C	Cor	E	H
Zone I1	2	47,44	29,51	47,83	34,56	53,61	47,44	29,51	47,83	34,56	53,61	47,44	29,51	47,83	34,56	53,61	47,44	29,51	47,83	34,97	53,61
	3	47,87	30,47	49,27	35,9	54,37	47,87	30,47	49,27	35,9	54,37	47,87	30,47	49,27	36,9	54,37	48,92	30,47	49,27	39,9	54,37
	4	49,31	31,37	50,44	37,78	54,8	49,31	31,37	50,44	37,78	54,8	49,31	31,37	50,44	37,44	54,8					
Zone I2	12	56,3	46,19	64,56	52,48	60,26	56,3	46,19	64,56	52,48	58,99	56,3	46,19	63,9	54,72	60,57					
	14	57	48,33	64,76	54,39	61,23	57	48,33	64,76	54,39	60,96	57	48,33	65,44	60,92	63,8					
	18	59,09	50,96	66,71	60,35	64,08	59,09	50,96	66,71	60,35	64,1	59,09	50,87	67,47	69,74	64,74					
	20	63,08	54,53	68,09	63,41	64,11	63,08	54,53	68,09	63,41	64,31	63,08	52,27	68,54	73,31	65,42					
Zone I3	25	66,12	56,63	70,6	74	67,39	64,89	57,9	70,6	74,05	65,85	64,89	59,42	73,33	78,28	66,69					
	35	76,06	64,78	75,57	78,24	70,31	74,72	67,39	76,27	78,63	71,99	75,59	65,75	73,8	80,77	66,35					
	36	76,49	65,58	75,82	78,01	72,67	76,6	67,68	77,67	79,17	72,14	76,19	66,92	73,8	81,39	66,35					
	40	78,96	67,74	76,99	79,19	75,24	79,29	68,32	78,26	81,41	77,3										
	45	80,85	69,41	79,33	81,9	75,9	81,01	69,84	81,02	82,12	76,64										
	50	81,63	71,28	80,93	82,37	79,17	81,12	70,54	82,08	82,86	77,17										
	53	82,27	72,39	81,49	82,84	79,21	86,03	71,03	82,04	83,46	76,76										
Zone I4	60	82,74	73,84	82,82	83,99	81,14	85,08	73,1	82,33	84,32	76,76										
	70	86,95	74,67	83,07	85,04	87,28															
	75	86,81	75,71	82,68	85,39	87,08															
	80	87,28	76,06	81,9	86,05	86,56															

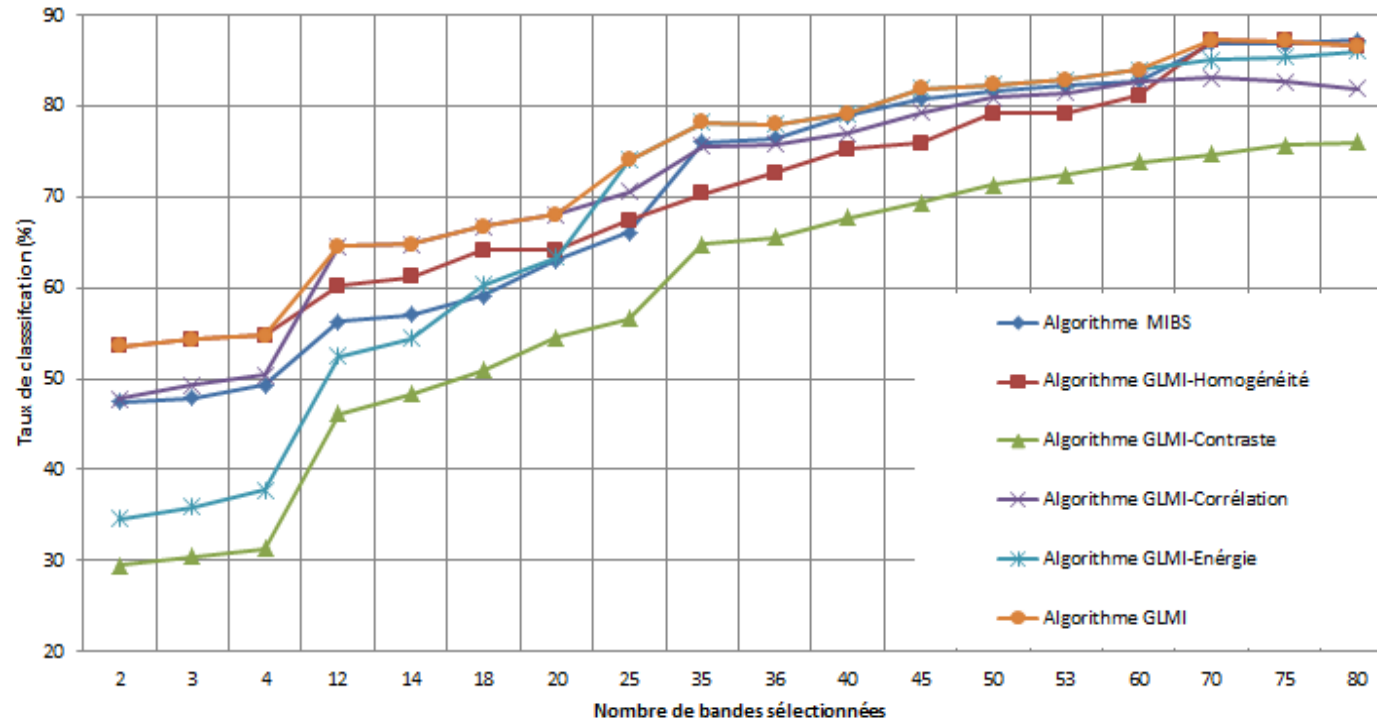


Figure 6. 8 : Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur AVIRIS Indian-Pines.

Tableau 6. 2: Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Salinas.

Le taux de classification (%) selon les différents seuils																						
		-0.02					-0.01					-0.005					0					
		MIBS	GLMI				MIBS	GLMI				MIBS	GLMI				MIBS	GLMI				
			C	Cor	E	H		C	Cor	E	H		C	Cor	E	H		C	Cor	E	H	
Zone S1	2	65,98	32,11	66,28	20,81	38,6	65,98	32,11	66,28	20,81	38,6	65,98	32,11	66,28	20,81	38,6	65,98	32,11	66,28	20,81	38,6	
	3	77,75	36,9	71,07	33,05	44,19	77,75	36,9	71,07	33,05	51,94	77,75	36,9	71,07	33,05	51,94	77,75	36,9	71,07	33,05	51,94	
	4	79,52	72,4	73,06	38,84	52,11	79,52	72,4	73,06	38,84	56,99	78,43	72,4	73,27	38,84	56,99						
Zone S2	12	83,55	85,51	85,29	73,95	76,83	83,55	85,51	84,9	75,19	77,16	80,81	85,51	85,12	75,19	77,37						
	14	85,23	86,2	86,2	76,63	81,38	85,23	86,2	85,33	79,59	83,89	80,81	86,2	85,38	79,59	84,03						
	18	86,25	87,12	86,4	85,16	83,61	86,25	86,54	85,7	85,7	86,15	80,81	86,54	86,39	85,87	86,67						
	20	86,51	87,59	87,15	85,99	86,69	86,51	86,55	87,19	86,31	86,73	80,81	86,54	86,39	86,32	87,71						
Zone S3	25	86,88	89,48	89,5	87,32	88,21	86,88	86,55	87,2	88,42	88,53	80,81	86,54	86,39	87,27	87,71						
	35	88,42	91,19	90,9	90,23	89,89	88,42	86,55	87,2	89,39	88,53	80,81	86,54	86,39	87,29	87,71						
	36	88,47	91,2	90,9	90,23	90,05	88,47	86,55	87,2	89,39	88,53	80,81	86,54	86,39	87,29	87,71						
Zone S4	40	88,97	91,22	91,92	90,23	91,03	88,51	86,55	87,2	89,39	88,53											
	45	91,11	91,22	94,59	90,23	91,37	88,54	86,55	87,2	89,39	88,54											
	50	91,14	91,22	94,81	90,23	91,37	88,54	86,55	87,2	89,39	88,54											
	53	91,77	91,22	94,87	90,23	91,37	88,54	86,55	87,2	89,39	88,54											
	60	91,85	91,22	95,16	90,23	91,37	88,54	86,55	87,2	89,39	88,54											
	70	92,86	91,22	95,43	90,23	91,37																
	75	93,11	91,22	95,63	90,23	91,37																
	80	93,19	91,22	95,7	90,23	91,37																

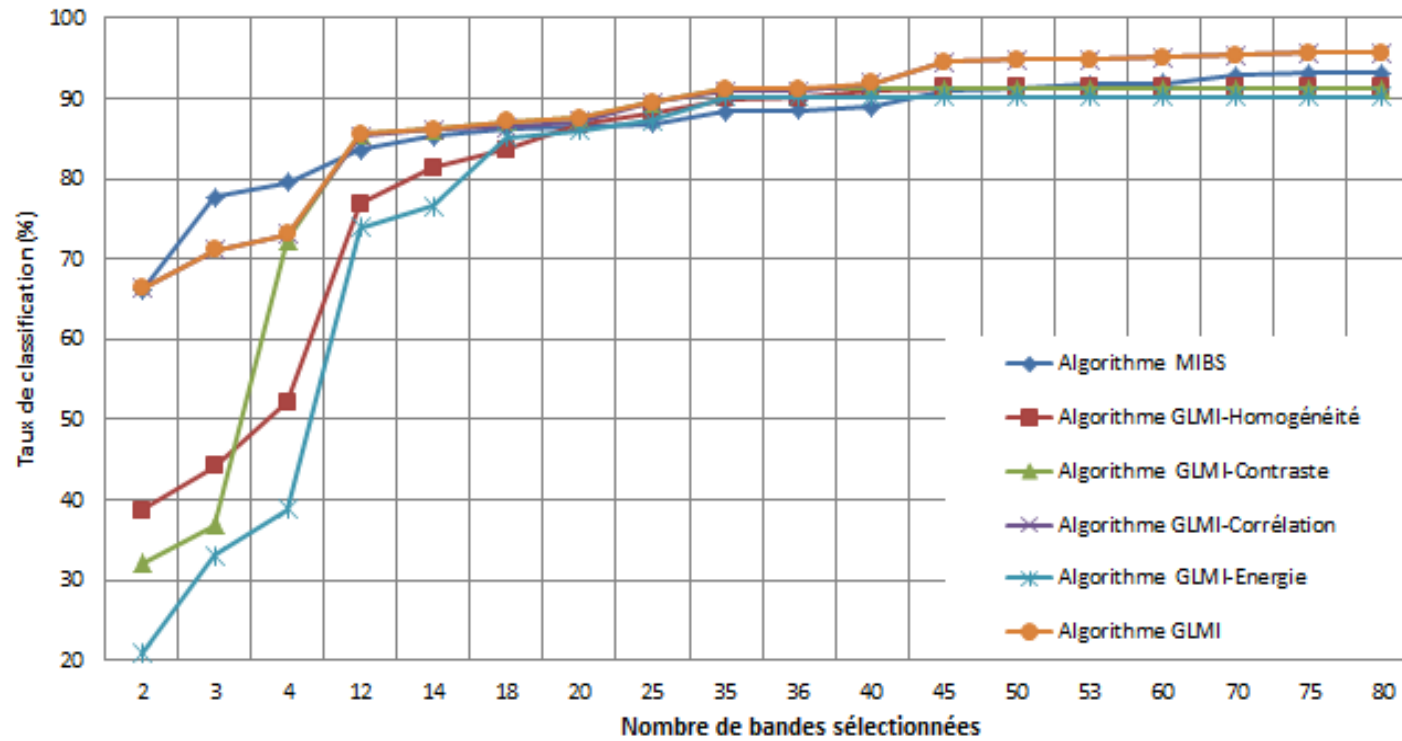


Figure 6. 9 : Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur AVIRIS Salinas.

Tableau 6. 3: Le taux de classification obtenu par l'algorithme reproduit MIBS et proposé GLMI en utilisant l'image Pavia-University.

Le taux de classification (%) selon les différents seuils																					
		-0.02					-0.01					-0.005					0				
		MIBS	GLMI				MIBS	GLMI				MIBS	GLMI				MIBS	GLMI			
			C	Cor	E	H		C	Cor	E	H		C	Cor	E	H		C	Cor	E	H
Zone P1	2	65,47	68,01	64,78	62,45	63,43	65,47	68,01	64,78	62,45	63,43	65,47	68,01	64,78	62,45	63,43	66,43	68,01	64,78	62,45	63,43
	3	66,61	75,55	65,27	62,61	63,87	66,61	75,55	65,27	62,61	63,87	66,61	75,55	65,27	62,61	63,87	66,43	75,55	65,27	63,4	64,85
	4	66,66	76,87	65,51	63,55	64,03	66,66	76,87	65,51	63,55	64,03	66,66	76,87	65,51	63,55	64,03					
	12	76,39	84,36	71,08	66,62	67,28	76,39	84,36	71,08	67,28	69,39	76,39	83,47	72,53	68,37	71,08					
	14	78,18	85,23	73,14	68,96	68,96	78,18	85,23	74,81	70,07	73,14	78,18	84,76	75,91	71,48	74,81					
	18	81,36	88,3	77,73	76,07	76,07	81,36	88,25	78,4	77,13	78,16	81,36	88,27	78,53	77,73	78,4					
	20	85,06	89,42	78,64	78,1	78,02	85,06	89,24	82,03	78,64	79,01	85,06	88,66	82,13	78,88	79,3					
	25	89,08	90,26	83,04	80,53	80,53	89,08	89,48	83,54	80,58	82,85	88,44	88,76	83,1	80,61	83,03					
Zone P2	35	91,64	90,27	84,46	87,36	85,32	89,76	89,48	85,63	84,25	85,09	88,44	88,76	85,04	84,16	84,21					
	36	91,64	90,27	85,52	87,66	85,32	89,76	89,48	85,72	84,27	85,11	88,44	88,76	85,13	84,34	84,53					
	40	91,64	90,27	85,81	88,11	85,62	89,76	89,48	86,21	85,38	85,45										
	45	91,64	90,27	87,25	89,29	86,86	89,76	89,48	89,54	88,03	89,09										
	50	91,64	90,27	89,7	89,5	89,38	89,76	89,48	89,7	89,06	89,6										
	53	91,64	90,27	89,7	90,09	89,7	89,76	89,48	89,73	89,76	90,21										
	60	91,64	90,27	90,62	91,01	91	89,76	89,48	90,59	90,56	90,86										
Zone P3	70	91,64	90,27	93,39	91,09	91,09															
	75	91,64	90,27	93,53	91,09	91,09															
	80	91,64	90,27	93,53	91,09	91,09															

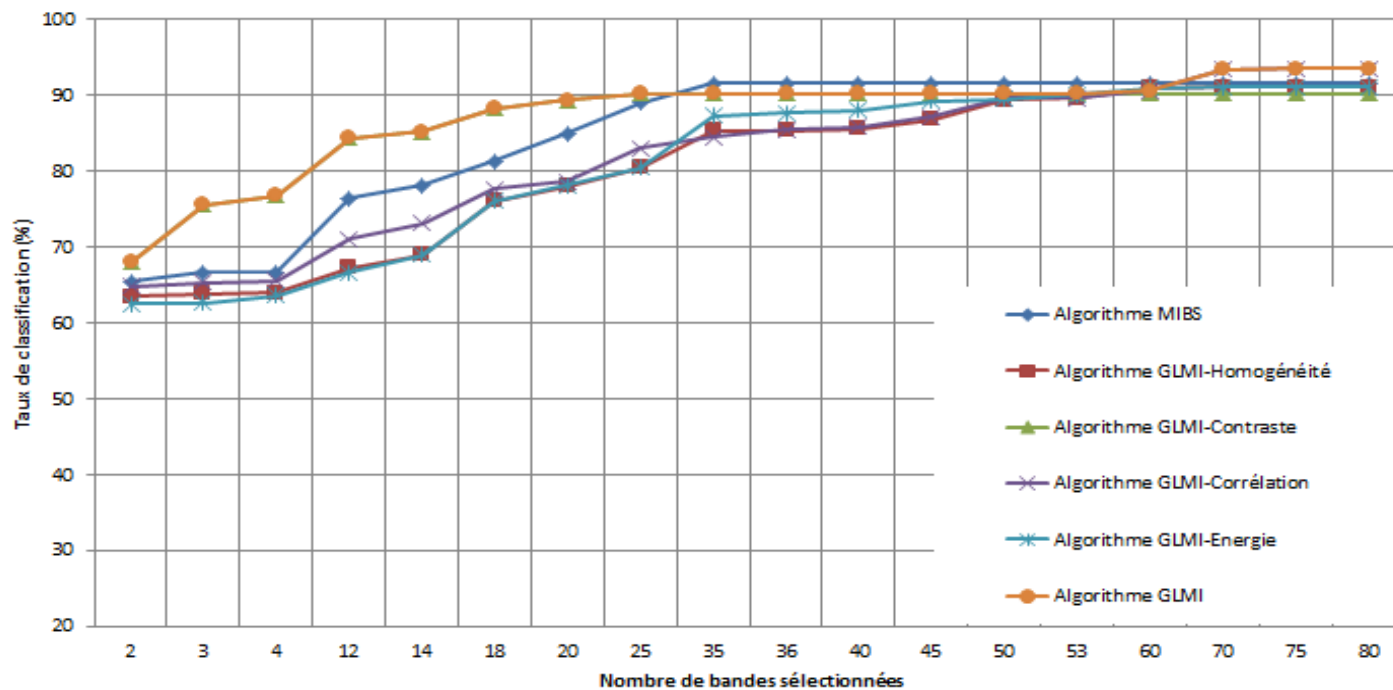


Figure 6. 10: Taux de la classification des méthodes MIBS et GLMI pour $Th=-0,02$ appliquées sur ROSIS_Pavia University

6.4.2.1.2 Analyse des résultats

Nous avons introduit le seuil T_h pour pouvoir contrôler la redondance permise. Pour $T_h=0$, la redondance entre les bandes sélectionnées est faible, peu de bandes sont sélectionnées. Pour des seuils négatifs, on admet plus de redondance ainsi, plus de bandes sont sélectionnées.

D'après les tableaux 6.1, 6. 2 et 6.3, et les figures 6.8, 6.9, 6.10, on remarque que :

- Pour l'IHS Indian-Pines,
 - Dans la zone I1, où l'algorithme GLMI doit choisir un nombre réduit de bandes, le critère d'homogénéité fournit des taux de classification élevés par rapport aux autres critères : Contraste, Corrélation et Energie. GLMI avec l'homogénéité est aussi plus performant que l'algorithme MIBS.
 - Dans la zone I2 où le nombre sélectionné est égal à une vingtaine de bandes, le critère de corrélation devient plus performant que les autres critères de GLMI et le MIBS.
 - Pour la zone I3 où le nombre de bandes sélectionnées atteint 60 bandes, le GLMI avec le critère d'Energie donne des résultats supérieurs au MIBS et les autres critères de GLMI.
 - Dans la zone I4, le seuil de contrôle de redondance de (-0.02) ne permet plus à l'algorithme de choisir des bandes pertinentes.
- Pour l'IHS Salinas,
 - Dans la zone S1, peu de bandes sont sélectionnées, le MIBS et GLMI avec la Corrélation donnent les bandes les plus performantes.
 - Pour la zone S2, une vingtaine de bandes est sélectionnée. Généralement, le GLMI avec le Contraste est plus performant.
 - Dans la zone S3 où le nombre de bandes sélectionnées atteint 60. Le Contraste et l'Énergie donnent les performances les plus élevées selon le seuil de la redondance.
 - Dans la zone S4, lorsqu'on permet plus de redondance, le GLMI-Contraste est plus performant.
- Pour l'IHS Pavia-University,
 - Lorsque la redondance admise est faible (Zone P1), le GLMI-Contraste donne des taux de classification supérieurs aux algorithmes MIBS, GLMI-Corrélation, GLMI-Energie et GLMI-Homogénéité.

- Si on admet plus de redondance jusqu'à 60 bandes sélectionnées (Zone P2), le MIBS devient plus performant que le GLMI. Mais au-delà de 60 bandes, le GLMI-Corrélation donne plus de précision de classification.
- Pour l'IHS Pavia-University, on remarque que le contraste améliore la sélection des bandes pertinentes. Par contre, le contraste n'est pas le bon critère pour classer la scène Indian-Pines. Il peut être considéré pour l'IHS Salinas.
- La corrélation (GLMI-Corrélation) permet une sélection plus précise que l'information mutuelle (MIBS).
- L'énergie s'apprête à une meilleure sélection lorsque les bandes sélectionnées sont nombreuses, la redondance est moyenne dans les deux cas de l'IHS Indian-Pines et Salinas.
- La sélection par homogénéité donne des résultats performants pour peu de bandes dans le cas d'Indian-Pines.

6.4.2.2 Résultats et analyse de la classification des classes des IHS selon les critères de GLMI : Contraste, Corrélation, Énergie et Homogénéité

Dans les expériences suivantes, nous calculons la classification individuelle (ICA) pour chaque classe des IHS utilisées avec l'algorithme proposé GLMI. Les résultats obtenus sont présentés dans le tableau 6.4 et figure 6.11 pour la scène Indian, dans le tableau 6.5 et figure 6.12 pour Salinas et dans le tableau 6.6 et figure 6.13 pour l'image Pavia-University. Les tableaux donnent l'ICA selon les critères Contraste, Corrélation, Énergie et Homogénéité. Le seuil de redondance T_h est fixé à -0.02. Le nombre de bandes X choisi est tel que le taux de classification des données de test T_x ne s'améliore plus considérablement en ajoutant plus de bandes sélectionnées.

Pour chaque IHS, les figures 6.11, 6.12 et 6.13 donnent la vérité de terrain et la scène reproduite en utilisant l'algorithme GLMI selon les critères: Contraste, Corrélation, Énergie et Homogénéité.

6.4.2.2.1 Résultats et analyse pour l'IHS Indian-Pines

La première colonne du tableau 6.4 représente le nombre total de pixel dans chaque classe d'Indian-Pines. Les autres colonnes montrent l'ICA obtenu en utilisant le sous-ensemble de bandes sélectionnées en fonction des différentes caractéristiques de texture extraites de la matrice GLCM (Contraste, Corrélation, Énergie et Homogénéité) en combinaison avec l'information mutuelle.

Tableau 6. 4: La précision de la classification individuelle ICA (%) de chaque classe de l'image Indian-Pines en utilisant l'algorithme proposé pour $Th=-0,02$

Classes	Nombre total des pixels	Précision de la classification individuelle ICA(%) pour le seuil $Th=-0,02$			
		Contraste X=40	Corrélation X=70	Energie X=80	Homogénéité X=70
1	54	39,13	78,26	86,96	82,61
2	1434	68,48	74,06	81,87	79,64
3	834	63,79	80,58	80,10	82,97
4	234	27,35	58,97	72,65	69,23
5	497	76,83	89,02	92,68	90,65
6	747	91,62	94,41	96,37	95,81
7	26	38,46	61,54	76,92	84,62
8	489	95,51	95,51	97,96	95,51
9	20	0,00	80,00	80,00	100,00
10	968	68,18	80,79	80,37	86,57
11	2468	82,01	82,25	86,47	88,09
12	614	64,82	81,11	85,34	87,30
13	212	94,17	98,06	98,06	98,06
14	1294	92,74	95,36	93,04	95,83
15	380	47,59	53,01	57,83	63,25
16	95	71,74	91,30	93,48	91,30

La Figures 6.11 donne la vérité de terrain et la scène reproduite en utilisant l'algorithme GLMI selon les critères: Contraste, Corrélation, Énergie et Homogénéité pour Indian-Pines.

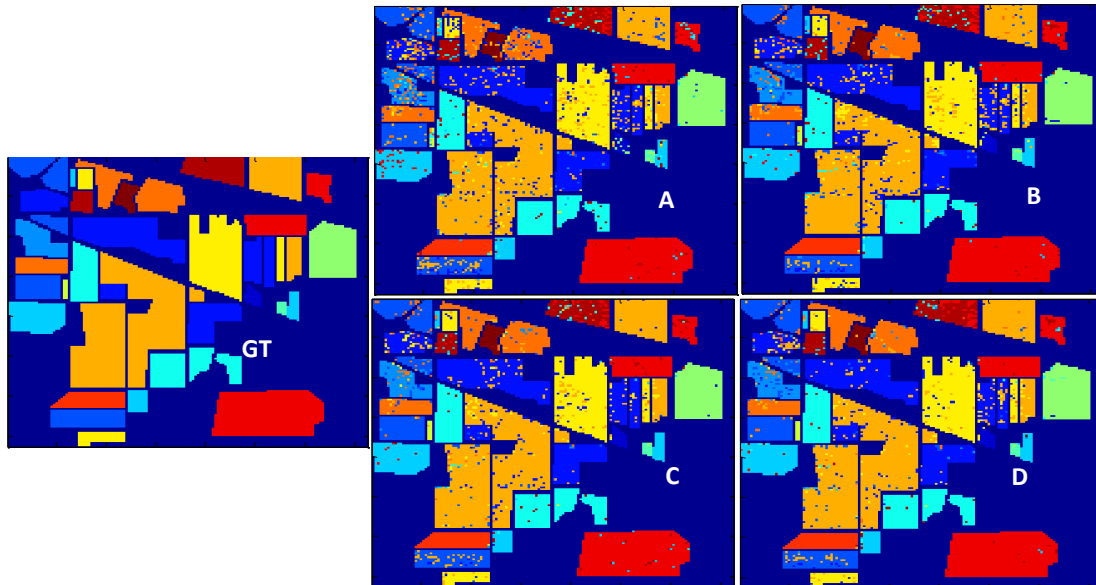


Figure 6. 11: La Vérité-terrain d'Indian-Pines (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D).

Analyse :

Selon le tableau 6.4, on peut remarquer que :

- L'homogénéité offre le meilleur potentiel pour distinguer les composantes des classes suivantes : Bois (classe 14 avec 95,83%), les trois classes Soja (10, 11 et 12) et la classe Avoine (classe 9 avec 100%) même si le nombre de pixels d'entraînement est minimum (seulement 10 pixels dans la dernière classe : c'est une classe homogène). Les résultats sont illustrés dans la figure 6.11 (D).
- L'énergie, par contre, donne les meilleurs résultats principalement pour les classes de (Luzerne:1), (Herbe: 5 et 6), et (Foins-andainé: 8) où les ICA sont respectivement : 86,96 %, 92,68 %, 96,37 %, 97,96 %, voir figure 6.11 (C).
- Il est aussi remarquable dans le tableau 6.4 que la corrélation donne la précision de classification maximale de 98,06 % pour la classe 13 (Blé), figure 6.11 (B).
- Le contraste classe mieux les classes (6, 13 et 14) que les autres classes.
- La classe 15 (Construction-herbe-arbre-engins) est relativement faiblement classifiée: son contenu est mixte.
- Les classes les plus homogènes sont mieux classifiées par l'Homogénéité l'énergie et la corrélation. Par contre, elles sont moins contrastées: la classe 6 (herbe/arbre) avec 95.81%, la classe 13 avec 98.06% et la classe 14 avec 95.83% et la classe 9 avec 100%.

6.4.2.2.2 Résultats et analyse pour l'IHS Salinas

La première colonne du tableau 6.5 représente le nombre total de pixels dans chaque classe de Salinas. Les autres colonnes montrent l'ICA obtenu en utilisant le sous-ensemble de bandes sélectionnées en fonction des différents critères de GLMI : Contraste, Corrélation, Énergie et Homogénéité.

Tableau 6. 5: La précision de la classification individuelle ICA (%) de chaque classe de l'image Salinas en utilisant l'algorithme proposé pour $Th=-0,02$

Classes	Nombre total des pixels	Précision de la classification individuelle ICA(%) pour le seuil $Th=-0,02$			
		Contraste X=40	Corrélation X=80	Energie X=35	Homogénéité X=45
1	2009	99,01	100	99,01	99,3
2	3726	99,35	100	99,84	99,89
3	1976	98,48	99,7	97,37	99,39
4	1394	99,26	99,56	99,71	99,85
5	2678	98,65	99,92	99,17	100
6	3959	99,6	99,9	99,95	99,95
7	3579	99,5	99,78	99,5	99,89
8	11271	87,46	92,42	84,55	86,18
9	6203	97,2	99,94	99,77	99,9
10	3278	94,94	98,29	94,63	96,89
11	1068	85,39	99,81	95,32	99,06
12	1927	98,03	99,9	99,48	99,9
13	916	98,25	98,91	98,69	98,47
14	1070	97,38	98,69	96,07	97,57
15	7268	69,94	81,33	57,59	62,76
16	1807	99,22	99,89	98	99,22

La figure 6.12 présente la vérité de terrain et la scène reproduite en utilisant l'algorithme GLMI selon les critères: Contraste, Corrélation, Énergie et Homogénéité pour l'IHS Salinas.

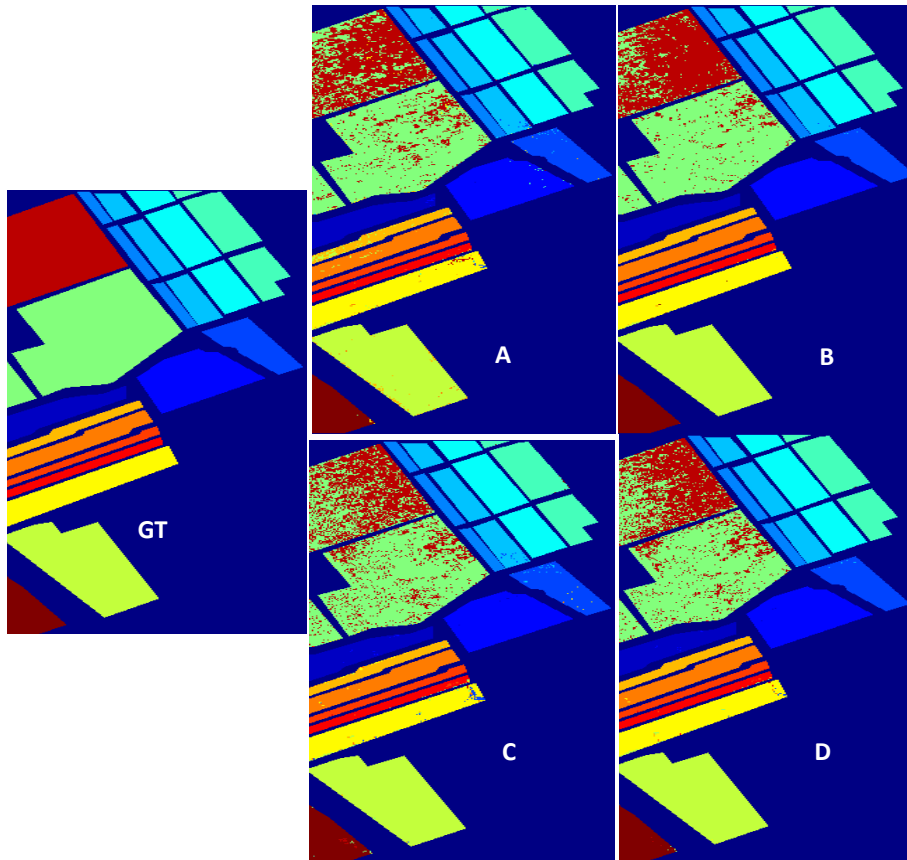


Figure 6. 12: La Vérité-terrain de Salinas (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D)

Analyse :

Pour les données d'AVIRIS_Salinas, à partir du tableau 6.5, nous pouvons faire les remarques suivantes :

- Comme dans le cas d'Indian-Pines, le critère de l'homogénéité fournit le meilleur ICA pour plusieurs classes : (Jachère : 4 et 5 avec 99,85% et 100%) et (Céliéri: 7 avec 99,89%), figure 6.12 (D).
- On voit aussi que l'énergie donne la précision individuelle maximale de classification de 99,95 % pour la classe 6 (Chaume), voir figure 6.12 (C).
- La corrélation, en revanche, donne les meilleurs résultats principalement pour toutes les classes Laitue (11, 12, 13 et 14 avec respectivement 99,81%, 99,9%, 98,91% et 98,69%), les classes de brocoli (1 et 2) sont également parfaitement classées en utilisant la corrélation avec ICA=100%. Les classes de Vigne' (15 et 16) avec respectivement (81,33% et 99,89%). L'ICA le plus élevé pour la classe 9 (sol de vigne) de 99,94 % est également obtenu en utilisant la caractéristique de corrélation, voir figure 6.12 (B).
- Le contraste donne des performances élevées de classification pour la majorité des classes, mais il reste inférieur aux autres critères.

6.4.2.2.3 Résultats et analyse pour l'IHS Pavia-University

Le tableau 6.6 représente l'ICA selon les classes obtenu en utilisant le sous-ensemble de bandes sélectionnées en fonction des différents critères de GLMI : Contraste, Corrélation, Énergie et Homogénéité pour l'IHS Pavia-University.

Tableau 6. 6: *La précision de la classification individuelle ICA (%) de chaque classe de l'image Pavia en utilisant l'algorithme proposé pour Th=-0,02*

Classes	Nombre total des pixels	Précision de la classification individuelle ICA(%) pour le seuil Th=-0,02			
		Contraste X=35	Corrélation X=75	Energie X=70	Homogénéité X=60
1	6631	90,65	95,45	94,6	94,54
2	18649	97,12	96,87	96,18	96,17
3	2099	64,72	78,59	71,51	70,55
4	3064	90,71	92,81	84,18	83,85
5	1345	100	99,85	99,85	99,85
6	5029	79,46	87,06	78,26	78,46
7	1330	72,58	87,27	85,91	85,61
8	3682	83,73	89,25	89,68	89,36
9	947	100	99,38	99,17	99,17

La figure 6.13 donne la vérité de terrain de l'IHS Pavia-University et la scène reproduite en utilisant l'algorithme GLMI selon les critères: Contraste, Corrélation, Énergie et Homogénéité.

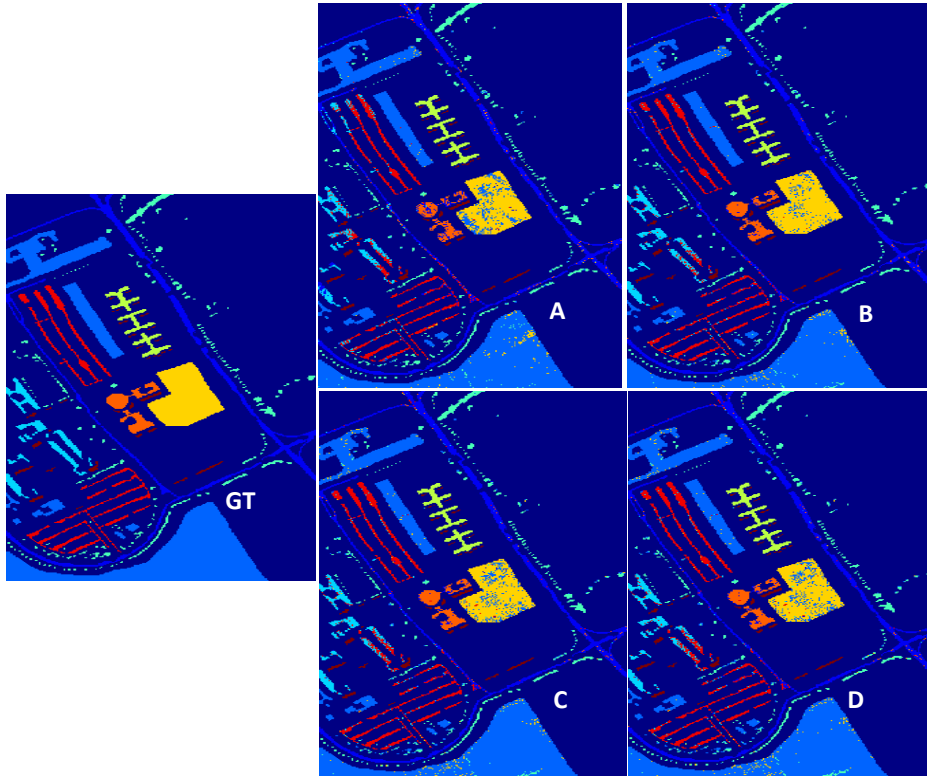


Figure 6. 13: La Vérité-terrain de Pavia (GT) et les cartes produites par notre algorithme proposé en utilisant le contraste (A), la corrélation (B), l'énergie (C) et l'homogénéité (D).

Analyse :

Pour l'image de ROSIS Pavia-University, selon le tableau 6.6, nous pouvons constater que :

- Selon les critères utilisés avec le GLMI, la corrélation donne le meilleur ICA principalement pour les classes (Arbres :4 avec 92,81%), (Gravier: 3 avec 78,59%) et comme dans le cas de Salinas, le sol nu (classe 6) est mieux classé en utilisant la corrélation avec 87,06%), voir figure 6.13 (B).
- Le contraste, en revanche, a la meilleure précision individuelle de classification principalement pour les classes suivantes : (Ombres: 9 avec 100%), (Prairies: 2 avec 97,12%) et (Métal peint: 5 à 100%), voir figure 6.13 (A). Notez que les dernières classes sont toutes de nature urbaine.

6.4.2.3 Résultats et analyse de la comparaison de l’algorithme proposé GLMI avec les algorithmes MIM, MIBS et DSEBS pour 64 bandes.

Dans cette section, nous introduisons la comparaison de la méthode proposées GLMI avec la méthode ‘Maximisation de l’information mutuelle MIM’ et la sélection des bandes par information mutuelle ‘MIBS’ [Viola, 1997 ; Sarhrouni, 2012]. Il s'agit des approches Filtres qui utilisent uniquement l'information spectrale pour la réduction de la dimensionnalité et la classification des IHS. Une autre méthode est prise en compte dans la comparaison qui utilise à la fois des informations spectrales et spatiales appelées ‘sélection non supervisée de bande hyperspectrale par l’ensemble dominant ‘Dominant-Set-Extraction-Based Selector DSEBS’ [Zhu, 2016] ; les étapes de cette méthode sont présentées en annexe. Les résultats de la classification SVM sans réduction de dimensionnalité sont également inclus dans la comparaison.

Le tableau 6.7 montre les résultats obtenus en utilisant l'algorithme proposé en comparaison avec les algorithmes MIM, MIBS et DSEBS pour les trois IHS Indian-Pines, Salinas et Pavia-University. Pour chacune des images hyperspectrales utilisées, les expériences ont été effectuées à l'aide des mêmes ensembles d'apprentissage et de test qui sont choisis arbitrairement, le rapport étant fixé à 1:10 pour chaque classe. La précision globale (OA) et le coefficient Kappa (k) sont utilisés comme critères d'évaluation des performances à cette étape. Le temps de calcul est également utilisé. La comparaison a été effectuée en utilisant 64 bandes sélectionnées à partir des trois images d'entrées.

Tableau 6. 7: La précision globale OA (%), le coefficient Kappa k et le temps d’exécution T(s) des trois images obtenus par les différentes méthodes en utilisant 64 bandes sélectionnées.

		SVM-toutes les bandes	MIM	MIBS	DSEBS	GLMI Proposé
Indian	OA(%)	60,74	71,67	78,36	81,01	81,86
	Kappa	0,5812	0,6978	0,7692	0,7975	0,8065
	Temps(s)	42,83	11,58	9,51	10,63	7,47
Salinas	OA(%)	87,31	88,75	89,73	90,36	94,24
	Kappa	0,8647	0,8800	0,8905	0,8972	0,9386
	Temps(s)	397,47	112,69	100,64	104,96	72,05
Pavia	OA(%)	78,44	82,45	84,83	85,09	90,07
	Kappa	0,7574	0,8025	0,8294	0,8322	0,8883
	Temps(s)	388	207,81	116,78	174,42	130,16

D'après le tableau 6.7, nous remarquons que:

- Les résultats de classification des quatre algorithmes (MIM, MIBS, DSEBS et GLMI) avec 64 bandes en utilisant les trois IHS sont supérieurs aux résultats du SVM en

utilisant toutes les bandes. Ceci est dû à la réduction de dimensionnalité effectuée par les quatre algorithmes : Ils éliminent les bandes non pertinentes ou redondantes qui diminuent la précision de la classification.

- Pour les trois scènes hyperspectrales, les deux méthodes spectrales-spatiales : GLMI et DSEBS améliorent significativement les précisions de classification par rapport aux MIM et MIBS qui sont basées uniquement sur des informations spectrales.
- La figure 6.14 illustre la précision globale OA et le coefficient kappa obtenus par les différentes méthodes sur les trois IHS utilisées. Il est clairement observé, à partir des bars d'OA, que la méthode proposée surpasse les autres méthodes avec 0,85 %, 3,5 %, 10,19 % et 21,21 % respectivement pour les méthodes DSEBS, MIBS, MIM et le SVM sans réduction de dimensionnalité lorsqu'elles sont appliquées aux données AVIRIS-Indian. Pour l'image AVIRIS-Salinas, aussi la méthode GLMI permet d'obtenir une meilleure précision globale OA qui dépasse DSEBS par 3,88%, MIBS par 4,51%, MIM de 5,49% et le SVM avec 6,93%. Dans le cas de ROSIS-Pavia, l'algorithme proposé fournit une amélioration du taux qui atteint 4,98%, 5,24%, 7,62% et 11,63% pour respectivement DSEBS, MIBS, MIM et SVM.

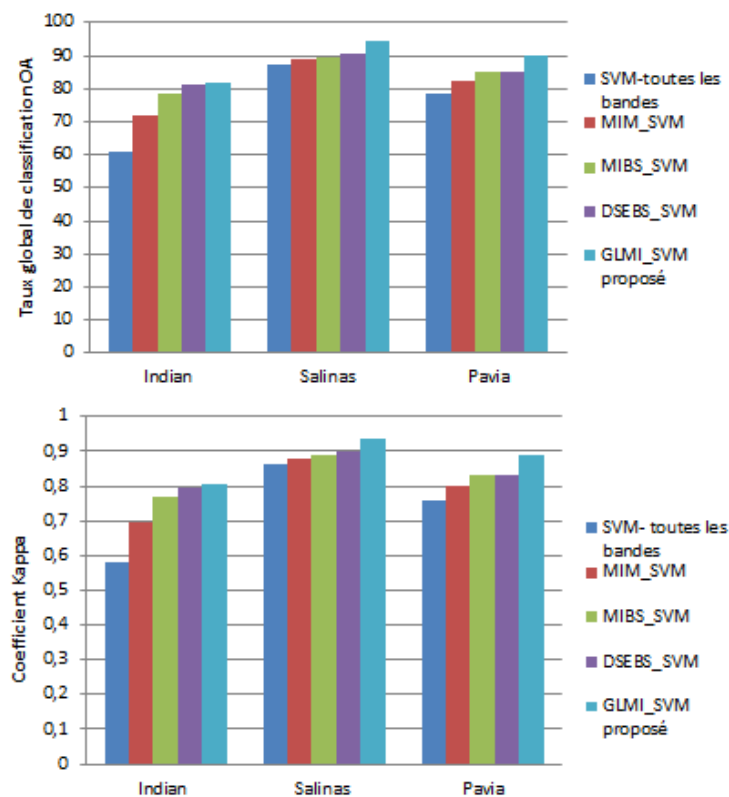


Figure 6. 14: Graphiques comparatifs des performances des trois images Indian, Salinas et Pavia en utilisant les différentes méthodes en termes d'OA et de Kappa.

- Les valeurs kappa illustrées dans la figure 6.14, permettent de faire les mêmes remarques que pour l'OA. L'algorithme GLMI domine les autres méthodes lorsqu'il est appliqué à Indian, Salinas et Pavia-University.
- En se référant au tableau 6.7, concernant le temps d'exécution de la réduction de dimensionnalité et de l'étape de classification, nous pouvons voir que le temps dépend de la taille des données utilisées. Généralement, l'intervalle de temps de simulation est compris entre : (7,47s et 42,83s) pour Indian-Pines, (72,05s et 397,47s) pour la scène de Salinas et (130,16s et 388s) pour celui de Pavia-University. On voit aussi que la classification sans réduction de dimensionnalité nécessite un temps considérable par rapport aux autres méthodes qui éliminent les bandes non pertinentes ou redondantes ce qui représente un avantage de la DR en termes de temps de calcul. Par rapport aux méthodes MIM, MIBS et DSEBS: le temps d'exécution du GLMI est généralement le plus faible.

6.5 Conclusion

La haute dimensionnalité des données hyperspectrales a imposé de nombreux problèmes dans les systèmes de traitement. Pour cela, la réduction de la dimensionnalité joue un rôle important avant la classification. Plusieurs travaux ont été développés dans ce domaine mais la problématique est toujours ouverte. Dans ce chapitre, nous avons proposé une nouvelle approche de type Filtre nommée GLMI. Elle combine des informations spectrales et spatiales pour réduire la dimensionnalité des IHS. Cela a permis d'améliorer les résultats de classification et réduire le temps d'exécution. La matrice GLCM a été retenue pour extraire les éléments de texture qui caractérisent l'information spatiale. Nous avons utilisé dans notre algorithme proposé : le contraste, la corrélation, l'énergie et l'homogénéité.

Nous avons appliqué l'algorithme proposé sur des données hyperspectrales de référence complexes capturées par le capteur AVIRIS (Indian-Pines et Salinas) et le capteur ROSIS (Pavia-University) en utilisant le classifieur SVM-RBF. Les résultats expérimentaux montrent l'efficacité de l'algorithme proposé qui procède à la sélection par la combinaison des caractéristiques spectrales (l'information mutuelle) et spatiales (Texture). La comparaison avec les méthodes MIM, MIBS et DSEBS montre que la méthode proposée GLMI permet d'augmenter la précision de la classification dans un temps de traitement réduit. Cette méthode est très intéressante à améliorer compte tenu de ses performances.

Conclusion générale

Dans ce mémoire de thèse, nous avons apporté de nouvelles approches pour contribuer à la résolution des problèmes de réduction de dimensionnalité et la classification des images hyperspectrales IHS. L'objectif des approches proposées est la sélection d'un groupe réduit des attributs pertinents avec contrôle de redondances entre les bandes sélectionnées. Ceci permettra d'augmenter les performances de la classification pour une meilleure discrimination des différents objets des scènes observées.

Premièrement, nous avons présenté le concept de l'imagerie hyperspectrale, ses principales caractéristiques, ses applications et ses modes d'acquisition. Ceci nous a permis de dégager les enjeux et problématiques liés à la classification de ces données de haute dimensionnalité. Deuxièmement, nous nous sommes intéressés à la problématique principale de la réduction de dimensionnalité des IHS. Nous avons dressé l'état de l'art des diverses procédures utilisées dans ce cadre. Les avantages et les limitations de chacune d'elles nous ont permis de choisir celles qui seront adoptées pour le développement de nos propres méthodes proposées dans ce mémoire.

Nous avons également évalué les performances de quatre classifieurs supervisés (SVM, KNN, RF et LDA) en termes de performances de classification sur les IHS. L'étude comparative a été conduite sur trois IHS de différentes caractéristiques en utilisant l'information mutuelle pour la réduction de leurs dimensionnalités. L'analyse des résultats obtenus nous a permis de retenir le classifieur SVM avec noyau RBF (SVM-RBF) pour la validation des algorithmes proposés dans ce travail de thèse.

Nous avons ensuite proposé une nouvelle approche Filtre, pour réduire la dimensionnalité des IHS, en se basant sur l'information mutuelle normalisée NMI. Elle consiste à sélectionner les bandes informatives et enlever celles qui sont redondantes. Ceci est réalisé par la maximisation de la NMI calculée pour chaque bande avec la vérité de terrain. Le groupe de bandes sélectionnées est validé par le classifieur SVM sur deux images hyperspectrales réelles du capteur AVIRIS de la NASA. La comparaison des résultats obtenus avec d'autres méthodes dans la littérature montre l'efficacité et la rapidité de l'approche proposée. Cependant, les performances de l'algorithme d'induction ne sont pas intégrées dans l'étape de génération des sous-ensembles de bandes.

Pour examiner l'inclusion du classifieur SVM dans le processus de génération des sous-ensembles de bandes pertinentes, nous avons proposé une nouvelle approche wrapper. Elle est basée sur la maximisation de l'information mutuelle normalisée et le contrôle de l'intervalle de la probabilité d'erreur entre la vérité de terrain et celle estimée par le classifieur SVM. L'application est faite sur les mêmes IHS du capteur AVIRIS. Les résultats obtenus sont meilleurs en comparaison avec l'approche filtre basée sur la NMI. Cependant, le temps d'exécution a augmenté. Cette approche sera bénéfique dans les applications où une grande précision de classification est exigée.

Une autre approche a été proposée pour évaluer la combinaison de l'information spectrale, exploitée dans les approches précédentes, avec l'information spatiale des bandes. Des caractéristiques de textures ont été extraites de chaque bande (Contraste, corrélation, énergie et homogénéité) pour analyser les relations inter-pixels. Les bandes sélectionnées en combinant l'information mutuelle et les caractéristiques de texture sont validées par le classifieur SVM. Les expériences ont été menées sur trois IHS de différentes caractéristiques. Les résultats obtenus montrent l'avantage de l'apport de l'information spatiale à la sélection des bandes pertinentes. Les performances de la classification s'améliorent dans un intervalle de temps réduit.

La réduction de la dimensionnalité des images hyperspectrales permet la production de cartes thématiques précises pour améliorer l'exploitation des ressources de la terre. Ce domaine de recherche reste encore très vaste. En termes de perspectives, plusieurs pistes peuvent être examinées :

- Proposition d'utiliser des techniques d'optimisation pour réduire le temps d'exécution de l'algorithme Wrapper basée sur l'utilisation de l'information mutuelle normalisée et le contrôle de l'intervalle de la probabilité d'erreur.
- Les résultats de l'approche basée sur l'utilisation conjointe de l'information spectrale-spatiale étaient très encourageants. Ce qui donne une future piste d'exploiter les relations inter-pixels par la texture ou autres caractéristiques.
- Les vérités de terrain contiennent un pourcentage important de pixels non libellés (AVIRIS_Indian Pines : 50%, AVIRIS_Salinas : 51% et ROSIS_Pavia University 79%). Ceci ouvre une voie très intéressante à investiguer : appliquer l'information mutuelle et l'analyse de texture avec des méthodes de classification non supervisée sur les IHS dans des approches filtres et wrapper.

Publications Scientifiques

1. [Nhaila, 2014] NHAILA, Hasna, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. A survey on fundamental concepts and practical challenges of Hyperspectral images. In : *Complex Systems (WCCS), 2014 Second World Conference on*. IEEE, 2014. p. 659-664.(indexé dans Scopus)
2. [Nhaila, 2018a] NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. Supervised classification methods applied to airborne hyperspectral images: Comparative study using mutual information. In: *Intelligent Computing in Data Sciences (ICDS), 2018 Second International Conference on*.Elsevier, 2018 (indexé dans Scopus).
3. [Nhaila, 2018b] NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. A Novel Filter Approach for Band Selection and Classification of Hyperspectral Remotely Sensed Images Using Normalized Mutual Information and Support Vector Machines. In : *International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning*. Springer, Cham, 2018. p. 521-530. (indexé dans Scopus)
4. [Nhaila, 2018c] NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. New wrapper method based on normalized mutual information for dimension reduction and classification of hyperspectral images. In: *Optimization and Applications (ICOA), 2018 4th International Conference on*. IEEE, 2018. p. 1-7. (indexé dans Scopus)
5. [Nhaila, 2015] NHAILA, Hasna, MERZOUQI, Maria, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. Hyperspectral images classification and Dimensionality Reduction using Homogeneity feature and mutual information. In: *Intelligent Systems and Computer Vision (ISCV), 2015*. IEEE, 2015. p. 1-5. (indexé dans Scopus)
6. [Nhaila, 2018d] NHAILA, Hasna, SARHROUNI, Elkebir and HAMMOUCH, Ahmed. A new filter for dimensionality reduction and classification of hyperspectral images using GLCM features and mutual information. *International Journal of Signal and Imaging Systems Engineering*, 2018, vol. 11, no 4, p. 193-205. (indexé dans Scopus)

Annexe

Les Algorithmes de comparaison

La méthode pertinence maximale et de redondance minimale MRMR

Peng [Peng, 2005] introduit une méthode de sélection incrémentale d'attributs basée sur un critère de dépendance statistique maximale en utilisant l'information mutuelle. Ce critère est appelé nous dérivons d'abord minimale-redondance-maximale-pertinence (MRMR)

L'algorithme de cette méthode est résumé comme suit :

- 1) Dépendance maximale MaxD:

En utilisant l'information mutuelle, sélectionner l'ensemble des attributs (bandes) S qui ont la plus grande dépendance avec la classe cible c (Vérité de terrain). Ceci est formulé comme suit:

$$\max D(B, c), D = IM(\{B_i, i = 1, \dots, m\}, c),$$

- 2) Choisir le premier attribut B_i qui satisfait maxD.

- 3) Processus de Sélection :

Répéter l'opération jusqu'à avoir le nombre de bandes à sélectionner : $|S|=X$

Pour chaque paire d'attributs, calculer $IM(B_i, B_s)$ avec $B_i \in B, B_s \in S$.

La sélection de l'attribut B_i doit maximiser la fonction d'évaluation suivante :

$$\text{MRMR} = \text{Argmax}(IM(B_i, c) - \frac{1}{S} \sum_{B_s \in S} I(B_i, B_s))$$

- 4) Sortie : groupe de bandes sélectionnées.

- 5) Fin
-

La méthode « Gray Wolf Optimizer GWO »

Medjahed [Medjahed, 2016] propose cette méthode basée sur l'optimisation pour réduire la dimensionnalité des images hyperspectrales. Il propose de reformuler le problème de la sélection de bande comme un problème combinatoire en modélisant une fonction objective basée sur des mesures de séparabilité de classe et des taux de classification en utilisant le « Gray Wolf Optimizer » [Mirjalili, 2014].

L'algorithme GWO est défini comme suit :

- 1) Définir la forme de la fonction objective proposée, généralement elle est la suivante :
 $J(b) = \text{taux de la classification} + \text{mesure de la séparabilité de la classe}$.
- 2) b est la variable à optimiser. C'est un vecteur binaire de taille identique au numéro de bande.
 Si $b_i = 1$ le numéro de bande i est sélectionné sinon la bande i n'est pas sélectionnée.
- 3) Définir cinq fonctions objectives $J_1(b)$, $J_2(b)$, $J_3(b)$, $J_4(b)$ et $J_5(b)$: tels que:

- La première fonction utilise seulement le taux de classification :

$$J_1(b) = J_a(b) = \frac{\sum_{i=1}^N \text{assess}(p_x)}{|N|}, n_i \in N, \text{ Avec la bande } b_i \text{ est sélectionnée } b_i=1$$

- La deuxième fonction objective utilise seulement la distance de Hausdroff :

$$J_2(b) = J_H(b) = \frac{\sum_{i=1}^N b_i \times D_{b_i}}{\sum_{i=1}^k D_{b_i}}$$

Avec : $\sum_{i=1}^N b_i \times D_{b_i}$ est la fonction de calcul de la distance de Hausdorff pour les bandes sélectionnées, et $\sum_{i=1}^k D_{b_i}$ est la distance de Hausdorff pour toutes les bandes.

- La troisième fonction utilise seulement la distance de Jeffries-Matusita :

$$J_3(b) = J_{JM}(b) = \frac{\sum_{i=1}^N b_i \times D_{b_i}}{\sum_{i=1}^k D_{b_i}}$$

Avec : $\sum_{i=1}^N b_i \times D_{b_i}$ est la fonction de calcul de la distance de Jeffries-Matusita pour les bandes sélectionnées, et $\sum_{i=1}^k D_{b_i}$ est la distance de Jeffries-Matusita pour toutes les bandes.

- La quatrième fonction est la somme du taux de classification et le terme Hausdroff :

$$J_4(b) = \omega_a \times J_a(b) + w_H \times J_H(b)$$

Avec : ω_a est la mesure du taux de classification et ω_H est la mesure de la distance de Hausdorff.

- La cinquième fonction est la somme du taux de classification et le terme JM:

$$J_5(b) = \omega_a \times J_a(b) + w_{JM} \times J_{JM}(b)$$

Avec : ω_H est la mesure de la distance de JM.

- 4) Une modification de l'algorithme GWO [Mirjalili, 2014] est utilisée pour l'adapter au problème de la sélection et ainsi optimiser les fonctions objectives.

- 5) Fin

La méthode « dominant-set-extraction-based selector DSEBS »

Zhu [Zhu, 2016] propose cette méthode pour la sélection des bandes hyperspectrales. Des mesures tenant compte la structure de l'information de la bande et de son indépendance sont effectuées. Puis, une formulation graphique de la sélection de la bande qui permet une recherche intégrée au moyen de l'extraction des ensembles dominants « Dominant Set Extraction ».

L'algorithme DSEBS est défini comme suit :

I. Mesure de la structure de la bande et son indépendance

1) Analyse de la consistance spatiale-Spectrale

- a) Générer la carte de gradients de l'image hyperspectrale H_l en utilisant trois gradients.
- b) Binarisation de $H_l \rightarrow B_l$.
- c) Définir le moniteur D_l qui compte les transitions (c'est-à-dire les changements de 0 à 1 et vice versa) de chaque pixel de ses voisins.

2) Mesure de l'information de la bande

- a) Déterminer le degré de structure dans une bande en utilisant la moyenne μ_{D_l} de D_l
- b) Mesurer l'information relative de chacune des bandes dans les voisinages locaux en utilisant noyau Gaussien où large θ_l indique une bande très informative :

$$\theta_l = \exp[-\lambda(\mu_{D_l} - \widetilde{\mu}_{D_l})]$$

Avec : $\widetilde{\mu}_{D_l}$ les résultats du filtre Gaussien de μ_{D_l} et $\lambda > 0$ contrôle la décroissance de la fonction Gaussienne.

- 3) Construire la matrice L qui contient les mesures de l'indépendance des bandes en utilisant la corrélation par paire de bandes. L'indépendance entre deux bandes k et l est défini comme:

$$l_{lk} = \exp(-\gamma d_l d_k^T)$$

II. Représentation graphique et optimization

- 1) les bandes spectrales à sélectionner sont représentées sous la forme d'un graphique :

$$G = (V, E, f, g)$$

Avec : $V = \{1, \dots, N\}$, N est le nombre de bandes spectrales ; $E \subseteq V \times V$

$$f(l) = \theta_l \text{ et } g(l, k) = l_{lk}$$

- 2) Définir l'optimum local de l'ensemble dominant

- a) Associer un vecteur binaire évalué en valeur réelle non négative. Une petite valeur indique un nœud faiblement associé à l'ensemble dominant, alors qu'une grande valeur indique une forte association ; les nœuds qui ne participent pas à l'ensemble dominant sont affectés de valeurs nulles.
- b) Utiliser la dynamique du réplicateur [Weibull, 1997] pour simplifier la recherche des solutions (locales).

III. Fin

Références

- ABLIN, R. and SULOCHANA, C. Helen. A survey of hyperspectral image classification in remote sensing. *International Journal of Advanced Research in Computer and Communication Engineering*, 2013, vol. 2, no 8, p. 2986-3000.
- ADÃO, Telmo, HRUŠKA, Jonáš, PÁDUA, Luís, BESSA, José, PERES, Emanuel, MORAIS, Raul, and SOUSA, Joaquim João. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing*, 2017, vol. 9, no 11, p. 1110.
- AGARWAL, Megha, and MAHESHWARI, R. P. Co-occurrence of maximal Haar-like wavelet filters for CBIR. *International Journal of Signal and Imaging Systems Engineering*, 2015, vol. 8, no 5, p. 316-330.
- AL AMRANI, Yassine, LAZAAR, Mohamed, and EL KADIRI, Kamal Eddine. Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis. *Procedia Computer Science*, 2018, vol. 127, p. 511-520.
- APLIN, P. and ATKINSON, Peter M. Sub-pixel land cover mapping for per-field classification. *International Journal of Remote Sensing*, 2001, vol. 22, no 14, p. 2853-2858.
- ARABLOUEI, Reza, GOAN, Ethan, GENSEMER, Stephen and KUSY, Branislav. Fast and robust pushbroom hyperspectral imaging via DMD-based scanning. In : *Novel Optical Systems Design and Optimization XIX*. International Society for Optics and Photonics, 2016. p. 99480A.
- ASPINALL, Richard J., MARCUS, W. Andrew, and BOARDMAN, Joseph W. Considerations in collecting, processing, and analysing high spatial resolution hyperspectral data for environmental investigations. *Journal of Geographical Systems*, 2002, vol. 4, no 1, p. 15-29.
- BALAKRISHNAMA, Suresh, and GANAPATHIRAJU, Aravind. Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, 1998, vol. 18, p. 1-8.
- BATTITI, Roberto. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on neural networks*, 1994, vol. 5, no 4, p. 537-550.
- BELGIU, Mariana, and DRĂGUȚ, Lucian. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, vol. 114, p. 24-31.
- BELLMAN, R. Curse of dimensionality. *Adaptive control processes: a guided tour*. Princeton, NJ, 1961.
- BERMEJO, Pablo, GÁMEZ, Jose A., and PUERTA, Jose M. Incremental wrapper-based subset selection with replacement: An advantageous alternative to sequential forward selection. In: *Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on*. IEEE, 2009. p. 367-374.

- BERMEJO, Pablo, GÁMEZ, José A., and PUERTA, José M. Speeding up incremental wrapper feature subset selection with Naive Bayes classifier. *Knowledge-Based Systems*, 2014, vol. 55, p. 140-147.
- BOUROUHO, A., JILBAB, A., NACIR, C., and HAMMOUCH, A. Comparison of classification methods to detect the Parkinson disease. In : *Electrical and Information Technologies (ICEIT), 2016 International Conference on*. IEEE, 2016. p. 421-424.
- BREIMAN, Leo. Random forests. *Machine learning*, 2001, vol. 45, no 1, p. 5-32.
- BRUNN, A., FISCHER, C., DITTMANN, C., and RICHTER, R. Quality assessment, atmospheric and geometric correction of airborne hyperspectral hmap data. In : *Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy*. 2003. p. 72-81.
- CAMPS-VALLS, Gustavo., and BRUZZONE, Lorenzo. Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, vol. 43, no 6, p. 1351-1362.
- CARIOU, Claude., and CHEHDI, Kacem. A new k-nearest neighbor density-based clustering method and its application to hyperspectral images. In : *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. IEEE, 2016. p. 6161-6164.
- CAWLEY, Gavin C., TALBOT, Nicola L., and GIROLAMI, Mark. Sparse multinomial logistic regression via bayesian l1 regularisation. In : *Advances in neural information processing systems*. 2007. p. 209-216.
- CEN, Haiyan, LU, Renfu, ZHU, Qibing., and MENDOZAET, Fernando. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biology and Technology*, 2016, vol. 111, p. 352-361.
- CHANG, Chein-I. (ed.). Hyperspectral data exploitation: theory and applications. *John Wiley & Sons*, 2007.
- CHANG, Chein-I., and LIU, Keng-Hao. Progressive band selection of spectral unmixing for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, vol. 52, no 4, p. 2002-2017.
- CHANG, Yang-Lang, HUANG, Min-Yu, CHEN, Silvia Wan-Ting., and AYELE, Amare-Anagaw. 2-Dimension cluster of multiple correlation coefficient with mutual information applied to hyperspectral image band selection. In : *2018 7th International Symposium on Next Generation Electronics (ISNE)*. IEEE, 2018. p. 1-4.
- CHEN, Chen, LI, Wei, SU, Hongjun., and LIU, Kui. Spectral-spatial classification of hyperspectral image based on kernel extreme learning machine. *Remote sensing*, 2014, vol. 6, no 6, p. 5795-5814.
- CHOMBOON, Kittipong, CHUJAI, Pasapichi, TEERARASSAMEE, Pongsakorn, KERDPRASOP, Kittisak., and KERDPRASOP, Nittaya. An empirical study of distance metrics for k-nearest neighbor algorithm. In : *Proceedings of the 3rd International Conference on Industrial Application Engineering*. 2015.

- COHEN, Jacob. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 1960, vol. 20, no 1, p. 37-46.
- CORDELL, Susan, QUESTAD, Erin J., ASNER, Gregory P., KINNEY, Kealoha M., THAXTON, Jarrod M., UOWOLO, Amanda, Brooks, Sam., and CHYNOWETH, Mark W. Remote sensing for restoration planning: how the big picture can inform stakeholders. *Restoration Ecology*, 2017, vol. 25, p. S147-S154.
- CORNUÉJOLS, Antoine, MICLET, Laurent, and KODRATOV, Y. Apprentissage artificiel, Concepts et algorithmes. Eyrolles, 2002. *Site Web du livre* <http://www.eyrolles.com/Sciences/Livre/apprentissage-artificiel-9782212110203>.
- COURBOT, Jean-Baptiste, MAZET, Vincent, MONFRINI, Emmanuel., and COLLET, Christophe. Extended faint source detection in astronomical hyperspectral images. *Signal Processing*, 2017, vol. 135, p. 274-283.
- COVER, Thomas M. and THOMAS, Joy A. *Elements of information theory*. John Wiley & Sons, 2012.
- COVER, Thomas M. and THOMAS, Joy A. Entropy, relative entropy and mutual information. *Elements of information theory*, 1991, vol. 2, p. 1-55.
- CRACKNELL, A. P. Review article Synergy in remote sensing-what's in a pixel?. *International Journal of Remote Sensing*, 1998, vol. 19, no 11, p. 2025-2047.
- DAI, Qiong, CHENG, Jun-Hu, SUN, Da-Wen., and ZENG, Xin-An. Advances in feature selection methods for hyperspectral image processing in food industry applications: a review. *Critical reviews in food science and nutrition*, 2015, vol. 55, no 10, p. 1368-1382.
- DANIELIS, Alessandro, GIORGI, Daniela, LARSSON, Marcus, STRÖMBERGB, Tomas, COLANTONIOA, Sara., and SALVETTIA, Ovidio. Lip segmentation based on Lambertian shadings and morphological operators for hyper-spectral images. *Pattern Recognition*, 2017, vol. 63, p. 355-370.
- DASH, Manoranjan., and LIU, Huan. Feature selection for classification. *Intelligent data analysis*, 1997, vol. 1, no 3, p. 131-156.
- DEAN, A. M., and SMITH, G. M. An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, 2003, vol. 24, no 14, p. 2905-2920.
- DELALIEUX, Stephanie, VAN AARDT, J. A. N., KEULEMANS, Wannes, SCHREVEENS, Eddie., and COPPIN, Pol. Detection of biotic stress (*Venturia inaequalis*) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications. *European Journal of Agronomy*, 2007, vol. 27, no 1, p. 130-143.
- DONG, Yanni, DU, Bo, ZHANG, Liangpei., and ZHANG, Lefei. Exploring locally adaptive dimensionality reduction for hyperspectral image classification: A maximum margin metric learning aspect. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, vol. 10, no 3, p. 1136-1150.

- DONOHU, David L. High-dimensional data analysis: The curses and blessings of dimensionality. *AMS math challenges lecture*, 2000, vol. 1, no 2000, p. 32.
- DUDA, Richard O., HART, Peter E., and STORK, David G. *Pattern classification*. John Wiley & Sons, 2012
- DZIEWIERZ, Jerzy, MCGINTY, John, MACFIONNGHAILE, Pol, et al. Applying hyperspectral imaging to continuous processing of pharmaceuticals. *Solving problems with spectral imaging*, 2016.
- ECHES, Olivier. *Méthodes Bayésiennes pour le démixage d'images hyperspectrales*. 2010. Thèse de doctorat. Institut National Polytechnique de Toulouse.
- ELMASRY, Gamal M., and NAKAUCHI, Shigeki. Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality—A comprehensive review. *Biosystems Engineering*, 2016, vol. 142, p. 53-82.
- ESTÉVEZ, Pablo A., TESMER, Michel, PEREZ, Claudio A., and ZURADA, Jacek M. Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, 2009, vol. 20, no 2, p. 189-201.
- FAN, Jianchao, HAN, Min., and WANG, Jun. Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation. *Pattern Recognition*, 2009, vol. 42, no 11, p. 2527-2540.
- FANO, Robert M., and HAWKINS, David. Transmission of information: A statistical theory of communications. *American Journal of Physics*, 1961, vol. 29, p. 793-794.
- FENG, Chao-Hui, MAKINO, Yoshio, OSHITA, Seiichi., and MARTÍN, Juan-Francisco-García. Hyperspectral imaging and multispectral imaging as the novel techniques for detecting defects in raw and processed meat products: current state-of-the-art research advances. *Food Control*, 2018, vol. 84, p. 165-176.
- FENG, Jie, JIAO, Licheng, LIU, Fang, SUN, Tao., and ZHANG, Xiangrong. Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low redundancy. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, vol. 53, no 5, p. 2956-2969.
- FERNANDEZ, Daniel, GONZALEZ, Carlos, MOZOS, Daniel., and LOPEZ, Sebastian. FPGA implementation of the principal component analysis algorithm for dimensionality reduction of hyperspectral images. *Journal of Real-Time Image Processing*, 2016, p. 1-12.
- FISHER, Peter. The pixel: a snare and a delusion. *International Journal of Remote Sensing*, 1997, vol. 18, no 3, p. 679-685.
- FISHER, Ronald A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 1936, vol. 7, no 2, p. 179-188.
- FOODY, G. M. Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution. *International Journal of Remote Sensing*, 1998, vol. 19, no 13, p. 2593-2599.

- FROHN, Robert C., and LOPEZ, Ricardo D. *Remote Sensing for Landscape Ecology: New Metric Indicators: Monitoring, Modeling, and Assessment of Ecosystems*. CRC Press, 2017.
- GALKOWSKI, Tomasz, STARCZEWSKI, Artur., and FU, Xiuju. Improvement of the multiple-view learning based on the self-organizing maps. In : *International Conference on Artificial Intelligence and Soft Computing*. Springer, Cham, 2015. p. 3-12.
- GALLANT, Stephen I. Connectionist expert systems. *Communications of the ACM*, 1988, vol. 31, no 2, p. 152-169.
- GAO, Bo-Cai., and GOETZ, Alexander FH. Column atmospheric water vapor and vegetation liquid water retrievals from airborne imaging spectrometer data. *Journal of Geophysical Research: Atmospheres*, 1990, vol. 95, no D4, p. 3549-3564.
- GHEYAS, Iffat A., and SMITH, Leslie S. Feature subset selection in large dimensionality domains. *Pattern recognition*, 2010, vol. 43, no 1, p. 5-13.
- GONÇALVES, Miguel, SILVA, Jose Silvestre., and BIOUCAS-DIAS, Jose. Classification of vegetation types in military region. In : *Electro-Optical Remote Sensing, Photonic Technologies, and Applications IX*. International Society for Optics and Photonics, 2015. p. 96490A.
- GONZALEZ, Rafael C., and WOODS, Richard E. *Digital image processing second edition*. Beijing: Publishing House of Electronics Industry, 2002, vol. 455.
- GOWEN, Aoife A., BURGER, J., O'CALLAGHAN, D., and O'DONNELL, C.P. Potential applications of hyperspectral imaging for quality control in dairy foods. In : *1st international workshop on computer image analysis in agriculture, Potsdam, Germany*. 2009.
- GOWEN, Aoife A., FENG, Yaoze, GASTON, Edurne., and VALDRAMIDIS, Vasilis. Recent applications of hyperspectral imaging in microbiology. *Talanta*, 2015, vol. 137, p. 43-54.
- GUO, Baofeng, DAMPER, Robert I., GUNN, Steve R., and NELSON, J.D.B. A fast separability-based feature-selection method for high-dimensional remotely sensed image classification. *Pattern Recognition*, 2008, vol. 41, no 5, p. 1653-1662.
- GUO, Baofeng, GUNN, Steve R., DAMPER, Robert I., and NELSON, J.D.B. Band selection for hyperspectral image classification using mutual information. *IEEE Geoscience and Remote Sensing Letters*, 2006, vol. 3, no 4, p. 522-526.
- GUYON, Isabelle., and ELISSEEFF, André. An introduction to variable and feature selection. *Journal of machine learning research*, 2003, vol. 3, no Mar, p. 1157-1182.
- HARALICK, Robert M., SHANMUGAM, Karthikeyan., and DINSTEN, Its'hak. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 1973, no 6, p. 610-621.
- HASANLOU, Mahdi., and SAMADZADEGAN, Farhad. Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier. *IEEE Geoscience and Remote Sensing Letters*, 2012, vol. 9, no 6, p. 1046-1050.

- HEGE, E. Keith, O'CONNELL, Dan, JOHNSON, William, BASTY, Shridhar., and DERENIAK, Eustace. Hyperspectral imaging for astronomy and space surveillance. In : *Imaging Spectrometry IX. International Society for Optics and Photonics*, 2004. p. 380-392.
- HELLMAN, Martin., and RAVIV, Josef. Probability of error, equivocation, and the Chernoff bound. *IEEE Transactions on Information Theory*, 1970, vol. 16, no 4, p. 368-372.
- HORIBE, Yasuichi. Entropy and correlation. *IEEE transactions on systems, man, and cybernetics*, 1985, no 5, p. 641-642
- HSU, Chih-Wei., and LIN, Chih-Jen. A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, 2002, vol. 13, no 2, p. 415-425.
- HUANG, Cheng-Lung., and WANG, Chieh-Jen. A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with applications*, 2006, vol. 31, no 2, p. 231-240.
- HUANG, Hong., and YANG, Mei. Dimensionality reduction of hyperspectral images with sparse discriminant embedding. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, vol. 53, no 9, p. 5160-5169.
- HUANG, Jinjie, CAI, Yunze., and XU, Xiaoming. A wrapper for feature selection based on mutual information. In : *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on. IEEE*, 2006. p. 618-621.
- HUANG, Kunshan, LI, Shutao, KANG, Xudong., and FANG, Leyuan. Spectral–spatial hyperspectral image classification based on KNN. *Sensing and Imaging*, 2016, vol. 17, no 1, p. 1.
- HUGHES, Gordon. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 1968, vol. 14, no 1, p. 55-63.
- HUNG, Chih-Cheng, KULKARNI, Sameer., and KUO, Bor-Chen. A new weighted fuzzy C-means clustering algorithm for remotely sensed image classification. *IEEE Journal of Selected Topics in Signal Processing*, 2011, vol. 5, no 3, p. 543-553.
- HUNG, Ming-Chih., and RIDD, Merrill K. A subpixel classifier for urban land-cover mapping based on a maximum-likelihood approach and expert system rules. *Photogrammetric Engineering and Remote Sensing*, 2002, vol. 68, no 11.
- JAIN, Deepak Kumar, DUBEY, Surendra Bilouhan, CHOUBEY, Rishin Kumar, SINHAL, Amit, ARJARIA, Siddharth Kumar, JAIN, Amar., and WANG, Haoxiang. An approach for hyperspectral image classification by optimizing SVM using self-organizing map. *Journal of Computational Science*, 2018, vol. 25, p. 252-259.
- JANECEK, Andreas, GANSTERER, Wilfried, DEMEL, Michael. And ECKER, Gerhard. On the relationship between feature selection and classification accuracy. In : *New Challenges for Feature Selection in Data Mining and Knowledge Discovery*. 2008. p. 90-105.
- JANSSEN, Lucas LF., and MOLENAAR, Martien. Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 1995, vol. 33, no 3, p. 749-758.

- JAY, Sylvain, GUILLAUME, Mireille, MINGHELLI, Audrey, DEVILLE, Yannick, CHAMID, Malik, LAFRANCE, Bruno., and SERFATY, Véronique. Hyperspectral remote sensing of shallow waters: Considering environmental noise and bottom intra-class variability for modeling and inversion of water reflectance. *Remote Sensing of Environment*, 2017, vol. 200, p. 352-367.
- JIA, Jianhua, YANG, Ning, ZHANG, Chao, YUE, Anzhi, YANG, Jianyu., and ZHU, Dehai. Object-oriented feature selection of high spatial resolution images using an improved Relief algorithm. *Mathematical and Computer Modelling*, 2013, vol. 58, no 3-4, p. 619-626.
- JIA, Sen, QIAN, Yuntao, LI, Jiming, LIU, Weixiang., and JI, Zhen. Feature extraction and selection hybrid algorithm for hyperspectral imagery classification. In : *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*. IEEE, 2010. p. 72-75.
- JIA, Sen, XIE, Yao, SHEN, Linlin., and DENG, Lin. Hyperspectral image classification using Fisher criterion-based Gabor cube selection and multi-task joint sparse representation. In : *2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2015. p. 1-4.
- KANG, Xudong, LI, Shutao, FANG, Leyuan., and BENEDIKTSSON, Jón Atli. Intrinsic image decomposition for feature extraction of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, vol. 53, no 4, p. 2241-2253.
- KESHAVA, Nirmal., and BOETTCHER, Peter W. Relationships between physical phenomena, distance metrics, and best-bands selection in hyperspectral processing. In : *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*. International Society for Optics and Photonics, 2001. p. 55-68.
- KHARROUBI, Jamal. *Etude de techniques de classement" Machines à vecteurs supports" pour la vérification automatique du locuteur*. 2002. Thèse de doctorat. Télécom ParisTech.
- KOHAVI, Ron., and JOHN, George H. Wrappers for feature subset selection. *Artificial intelligence*, 1997, vol. 97, no 1-2, p. 273-324.
- KOONSANIT, Kitti, JARUSKULCHAI, Chuleerat., and EIUMNOH, Apisit. Band selection for dimension reduction in hyper spectral image using integrated information gain and principal components analysis technique. *International Journal of Machine Learning and Computing*, 2012, vol. 2, no 3, p. 248.
- KUNCHEVA, Ludmila. *Fuzzy classifier design*. Springer Science & Business Media, 2000.
- KVALSETH, Tarald O. Entropy and correlation: Some comments. *IEEE Transactions on Systems, Man, and Cybernetics*, 1987, vol. 17, no 3, p. 517-519.
- KWAK, Nojun., and CHOI, Chong-Ho. Input feature selection by mutual information based on Parzen window. *IEEE transactions on pattern analysis and machine intelligence*, 2002, vol. 24, no 12, p. 1667-1671.
- KWAK, Nojun. Feature extraction based on direct calculation of mutual information. *International Journal of Pattern Recognition and Artificial Intelligence*, 2007, vol. 21, no 07, p. 1213-1231.

- LAURIN, Gaia Vaglio, PULETTI, Nicola, HAWTHORNE, William, LIESENBERG, Veraldo, CORONA, Piermaria, PAPAIE, Dario, CHEN, Qi., and VALENTINI, Riccardo. Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sensing of Environment*, 2016, vol. 176, p. 163-176.
- LEUNG, K. Ming. Naive bayesian classifier. *Polytechnic University Department of Computer Science/Finance and Risk Engineering*, 2007.
- LI, IHSao-Chi, CHANG, Chein-I., and SONG, Meiping. Recursive Band Processing of Orthogonal Subspace Projection for Hyperspectral Imagery. *IEEE Geosci. Remote Sensing Lett.*, 2016, vol. 13, no 1, p. 3-7.
- LI, Wei, DU, Qian, ZHANG, Fan., and HU, Wei. Collaborative-representation-based nearest neighbor classifier for hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 2015, vol. 12, no 2, p. 389-393.
- LI, Wei, PRASAD, Saurabh., and FOWLER, James E. Hyperspectral image classification using Gaussian mixture models and Markov random fields. *IEEE Geoscience and Remote Sensing Letters*, 2014, vol. 11, no 1, p. 153-157.
- LI, Wei, PRASAD, Saurabh, FOWLER, James E., and BRUCE, Lori Mann. Locality-preserving dimensionality reduction and classification for hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, vol. 50, no 4, p. 1185-1198.
- LI, Yunsong, XIE, Weiyang., and LI, Huaqing. Hyperspectral image reconstruction by deep convolutional neural network for classification. *Pattern Recognition*, 2017, vol. 63, p. 371-383.
- LIN, Jianzhe, HE, Chen, WANG, Z. Jane., and LI, Shuying. Structure preserving transfer learning for unsupervised hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2017, vol. 14, no 10, p. 1656-1660.
- LIU, Dan, ZENG, Xin-An., and SUN, Da-Wen. Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: a review. *Critical reviews in food science and nutrition*, 2015, vol. 55, no 12, p. 1744-1757.
- LIU, Huan., and YU, Lei. Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering*, 2005, vol. 17, no 4, p. 491-502.
- LIU, Keng-Hao, CHIEN, Hung-Chang., and CHEN, Shih-Yu. Progressive sample processing of band selection for hyperspectral imagery. In : *Image and Signal Processing for Remote Sensing XXIII*. International Society for Optics and Photonics, 2017. p. 104271L.
- LIU, Yuwei, PU, Hongbin., and SUN, Da-Wen. Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications. *Trends in Food Science & Technology*, 2017, vol. 69, p. 25-35.
- LOH, Wei-Yin. Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2011, vol. 1, no 1, p. 14-23.

- LÖW, F., MICHEL, U., DECH, S., and CONRAD, C. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. *ISPRS journal of photogrammetry and remote sensing*, 2013, vol. 85, p. 102-119.
- LU, Dengsheng., and WENG, Qihao. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 2007, vol. 28, no 5, p. 823-870.
- LU, Guolan., and FEI, Baowei. Medical hyperspectral imaging: a review. *Journal of biomedical optics*, 2014, vol. 19, no 1, p. 010901.
- LUO, Yong, TAO, Dacheng, WEN, Yonggang, RAMAMOHANARAO, Kotagiri., and XU, Chao. Tensor canonical correlation analysis for multi-view dimension reduction. *IEEE transactions on Knowledge and Data Engineering*, 2015, vol. 27, no 11, p. 3111-3124.
- MA, Zhuang, LU, Yichao., and FOSTER, Dean. Finding linear structure in large datasets with scalable canonical correlation analysis. In : *International Conference on Machine Learning*. 2015. p. 169-178.
- MAKKI, Ihab, YOUNES, Rafic, FRANCIS, Clovis, BIANCHI, Tiziano., and ZUCCHETTI, Massimo. A survey of landmine detection using hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, vol. 124, p. 40-53.
- MAO, K. Z. Feature subset selection for support vector machines through discriminative function pruning analysis. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, vol. 34, no 1, p. 60-67.
- MARTÍNEZ, P., GUALTIERI, J. A., AGUILAR, P. L., PEREZ, R.M, LINAJE, M, PRECIADO, J.C., and PLAZA, A. Hyperspectral image classification using a self-organizing map. In : *Summaries of the X JPL Airborne Earth Science Workshop*. 2001.
- MEDJAHED, Seyyid Ahmed, SAADI, T. Ait, BENYETTOU, Abdelkader., and OUALI, Mohammed. Gray wolf optimizer for hyperspectral band selection. *Applied Soft Computing*, 2016, vol. 40, p. 178-186.
- MELESSE, Assefa M., and JORDAN, Jonathan D. A comparison of fuzzy vs. augmented-ISODATA classification algorithms for cloud-shadow discrimination from Landsat images. *Photogrammetric Engineering and Remote Sensing*, 2002, vol. 68, no 9, p. 905-912.
- MIRJALILI, Seyedali, MIRJALILI, Seyed Mohammad., and LEWIS, Andrew. Grey wolf optimizer. *Advances in engineering software*, 2014, vol. 69, p. 46-61.
- MOUNTRAKIS, Giorgos, IM, Jungho., and OGOLE, Caesar. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2011, vol. 66, no 3, p. 247-259.
- MUELLER, Andreas A., HAUSOLD, Andrea., and STROBL, Peter. HySens-DAIS/ROSIIS imaging spectrometers at DLR. In : *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology*. International Society for Optics and Photonics, 2002. p. 225-236.
- NAGANATHAN, Govindarajan Konda, GRIMES, Lauren M., SUBBIAH, Jeyamkondan, CALKINS, Chris R., SAMAL, Ashok., and MEYER, George E. Partial least squares analysis of near-

- infrared hyperspectral images for beef tenderness prediction. *Sensing and Instrumentation for Food Quality and Safety*, 2008, vol. 2, no 3, p. 178-188.
- NAGASUBRAMANIAN, Koushik, JONES, Sarah, SARKAR, Soumik, SINGH, Asheesh K., SINGH, Arti., and GANAPATHYSUBRAMANIAN, Baskar. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant methods*, 2018, vol. 14, no 1, p. 86.
- NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. A Novel Filter Approach for Band Selection and Classification of Hyperspectral Remotely Sensed Images Using Normalized Mutual Information and Support Vector Machines. In : *International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning*. Springer, Cham, 2018. p. 521-530.
- NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. New wrapper method based on normalized mutual information for dimension reduction and classification of hyperspectral images. In: *Optimization and Applications (ICOA), 2018 4th International Conference on*. IEEE, 2018. p. 1-7.
- NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. Supervised classification methods applied to airborne hyperspectral images: Comparative study using mutual information. In: *Intelligent Computing in Data Sciences (ICDS), 2018 Second International Conference on*. Elsevier, 2018.
- NHAILA, Hasna, MERZOUQI, Maria, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. Hyperspectral images classification and Dimensionality Reduction using Homogeneity feature and mutual information. In : *Intelligent Systems and Computer Vision (ISCV), 2015*. IEEE, 2015. p. 1-5.
- NHAILA, Hasna, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. A new filter for dimensionality reduction and classification of hyperspectral images using GLCM features and mutual information. *International Journal of Signal and Imaging Systems Engineering*, 2018, vol. 11, no 4, p. 193-205.
- NHAILA, Hasna, SARHROUNI, Elkebir., and HAMMOUCH, Ahmed. A survey on fundamental concepts and practical challenges of Hyperspectral images. In : *Complex Systems (WCCS), 2014 Second World Conference on*. IEEE, 2014. p. 659-664.
- NIAZMARDI, Saeid, HOMAYOUNI, Saeid., and SAFARI, Abdolreza. An improved FCM algorithm based on the SVDD for unsupervised hyperspectral data classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, vol. 6, no 2, p. 831-839.
- PAL, Mahesh., and FOODY, Giles M. Feature selection for classification of hyperspectral data by SVM. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, vol. 48, no 5, p. 2297-2307.
- PENG, Hanchuan, LONG, Fuhui., and DING, Chris. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 2005, vol. 27, no 8, p. 1226-1238.

- PICÓN, Ana J., VÁSQUEZ, Ramón, GONZÁLEZ, Jorge, LUVALL, Jeff., and RICKMAN, Douglas. Use of remote sensing observations to study the urban climate on tropical coastal cities. *Revista Umbral (Etapa IV-Colección completa)*, 2017, no 1, p. 218-232.
- PRASAD, S. V. S., SAVITHRI, T. Satya., and KRISHNA, Iyyanki V. Murali. Comparison of Accuracy Measures for RS Image Classification using SVM and ANN Classifiers. *International Journal of Electrical and Computer Engineering (IJECE)*, 2017, vol. 7, no 3, p. 1180-1187.
- QIAN, Yuntao, YE, Minchao., and ZHOU, Jun. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, vol. 51, no 4, p. 2276-2291.
- RAJADELL, Olga, GARCÍA-SEVILLA, Pedro., and PLA, Filiberto. Spectral–spatial pixel characterization using Gabor filters for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2013, vol. 10, no 4, p. 860-864.
- RASCHKA, Sebastian. 'Linear Discriminant Analysis bit by bit'. *Disponible en: sebastianraschka.com/Articles/2014_python_lda.html*, 2014.
- REN, Jinchang, KELMAN, Timothy., and MARSHALL, Stephen. Adaptive clustering of spectral components for band selection in hyperspectral imagery. In : *Hyperspectral Imaging Conference 2011*. 2011. p. 90-93.
- RIAZA, A., GARCIA MELENDEZ, E., SUAREZ, M., HAUSOLD, A., BEISL, U., and VAN DER WERFF, H. Mapping paleoflooded and emerged areas around lake marshes (Tablas de Daimiel, Spain) using hyperspectral DAIS 7915 and ROSIS Spectrometer data. In : Presented at the 3rd EARSeL Workshop on Imaging Spectroscopy. 2003.
- ROBNIK-ŠIKONJA, Marko., and KONONENKO, Igor. Theoretical and empirical analysis of ReliefF and RReliefF. *Machine learning*, 2003, vol. 53, no 1-2, p. 23-69.
- ROSCHER, Ribana., and WASKE, Björn. Shapelet-based sparse representation for land cover classification of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, vol. 54, no 3, p. 1623-1634.
- SAFAVIAN, S. Rasoul., and LANDGREBE, David. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 1991, vol. 21, no 3, p. 660-674.
- SAQUI, Diego, SAITO, José H., LUCIO, A. De C, FERREIRA, Ednaldo J, LIMA, Daniel C., and HERRERA, João P. Methodology for band selection of hyperspectral images using genetic algorithms and gaussian maximum likelihood classifier. In : *Computational Science and Computational Intelligence (CSCI), 2016 International Conference on*. IEEE, 2016. p. 733-738.
- SARHROUNI, ELkebir, HAMMOUCH, Ahmed., and ABOUTAJDINE, Driss. Dimensionality reduction and classification feature using mutual information applied to hyperspectral images: a filter strategy based algorithm. *arXiv preprint arXiv:1210.0052*, 2012.
- SARHROUNI, Elkebir, HAMMOUCH, Ahmed., and ABOUTAJDINE, Driss. Feature selection intelligent algorithm with mutual information and steepest ascent strategy. *International Journal of Advanced Intelligence Paradigms*, 2013, vol. 5, no 4, p. 257-269.
- SCHALKOFF, Robert J. *Artificial neural networks*. New York : McGraw-Hill, 1997.

- SCHMIDT, Johannes, FASSNACHT, Fabian Ewald, LAUSCH, Angela., and SCHMIDTLEIN, Sebastian. Assessing the functional signature of heathland landscapes via hyperspectral remote sensing. *Ecological indicators*, 2017, vol. 73, p. 505-512.
- SHAHSHAHANI, Behzad M., and LANDGREBE, David A. The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. *IEEE Transactions on Geoscience and remote sensing*, 1994, vol. 32, no 5, p. 1087-1095.
- SHAKHNAROVICH, Gregory. *Learning task-specific similarity*. 2005. Thèse de doctorat. Massachusetts Institute of Technology.
- SIGNORILE, Franck., and EMC-CAPTIVA, Invité. Extraction de séquences numériques dans des documents manuscrits quelconques Clément Chatelain. 2006.
- SONG, Benqin, LI, Jun, DALLA MURA, Mauro, LI, Peijun, PLAZA, Antonio, BIOUCAS-DIAS, José M, BENEDIKTSSON, Jon Atli., and CHANUSSOT, Jocelyn. Remotely sensed image classification using sparse representations of morphological attribute profiles. *IEEE transactions on geoscience and remote sensing*, 2014, vol. 52, no 8, p. 5122-5136.
- STREHL, Alexander., and GHOSH, Joydeep. Cluster ensembles---a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 2002, vol. 3, no Dec, p. 583-617.
- STUDHOLME, Colin, HILL, Derek LG., and HAWKES, David J. An overlap invariant entropy measure of 3D medical image alignment. *Pattern recognition*, 1999, vol. 32, no 1, p. 71-86.
- SUN, Kang, GENG, Xiurui, JI, Luyan., and LU, Yun. A new band selection method for hyperspectral image based on data quality. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, vol. 7, no 6, p. 2697-2703.
- TANG, Jiliang, ALELYANI, Salem., and LIU, Huan. Feature selection for classification: A review. *Data classification: Algorithms and applications*, 2014, p. 37.
- TSAI, Fuan., and LAI, Jhe-Syuan. Feature extraction of hyperspectral image cubes using three-dimensional gray-level cooccurrence. *IEEE transactions on geoscience and remote sensing*, 2013, vol. 51, no 6, p. 3504-3513.
- TYAGI, Mayank, BOVOLO, Francesca, MEHRA, Ankit K, CHAUDHURI, Subhasis., and BRUZZONE, Lorenzo. A context-sensitive clustering technique based on graph-cut initialization and expectation-maximization algorithm. *IEEE Geoscience and Remote Sensing Letters*, 2008, vol. 5, no 1, p. 21-25.
- UCHINO, Eiji, TOKUNAGA, Kazuhiro, TANAKA, Hiroki., and SUETAKE, Noriaki. IVUS-Based Coronary Plaque Tissue Characterization Using Weighted Multiple k-Nearest Neighbor. *Engineering Letters*, 2012, vol. 20, no 3.
- VAN DER MEER, Freek D., and DE JONG, Steven M. (ed.). *Imaging spectrometry: basic principles and prospective applications*. Springer Science & Business Media, 2011.
- VAPNIK, Vladimir. *The nature of statistical learning theory*. Springer science & business media, 1995.

- VIOLA, Paul., and WELLS III, William M. Alignment by maximization of mutual information. *International journal of computer vision*, 1997, vol. 24, no 2, p. 137-154.
- WAGSTAFF, Kiri, CARDIE, Claire, ROGERS, Seth., and SCHROEDL, Stefan. Constrained k-means clustering with background knowledge. In : *ICML*. 2001. p. 577-584.
- WEIBULL, Jörgen W. *Evolutionary game theory*. MIT press, 1997.
- WOODCOCK, Curtis E., and GOPAL, Sucharita. Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *International Journal of Geographical Information Science*, 2000, vol. 14, no 2, p. 153-172.
- WU, Hao., and PRASAD, Saurabh. Semi-supervised dimensionality reduction of hyperspectral imagery using pseudo-labels. *Pattern Recognition*, 2018, vol. 74, p. 212-224.
- XIA, Junshi, CHANUSSOT, Jocelyn, DU, Peijun., and HE, Xiyan. Rotation-based support vector machine ensemble in classification of hyperspectral data with limited training samples. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, vol. 54, no 3, p. 1519-1531.
- XIE, Li, LI, Guangyao, XIAO, Mang, PENG, Lei., and CHEN, Qiaochuan. Hyperspectral image classification using discrete space model and support vector machines. *IEEE Geoscience and Remote Sensing Letters*, 2017, vol. 14, no 3, p. 374-378.
- YANG, Chen, TAN, Yulei, BRUZZONE, Lorenzo, LU, Laijun., and GUAN, Renchu. Discriminative feature metric learning in the affinity propagation model for band selection in hyperspectral images. *Remote Sensing*, 2017, vol. 9, no 8, p. 782.
- YANG, Jinn-Min, YU, Pao-Ta., and KUO, Bor-Chen. A nonparametric feature extraction and its application to nearest neighbor classification for hyperspectral image data. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, vol. 48, no 3, p. 1279-1293.
- YU, Chunyan, LEE, Li-Chien, CHANG, Chein-I, XUE, Bai, SONG, Meiping., and CHEN, Jian. Band-Specified Virtual Dimensionality for Band Selection: An Orthogonal Subspace Projection Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, vol. 56, no 5, p. 2822-2832.
- YU, Enzhe., and CHO, Sungzoon. GA-SVM wrapper approach for feature subset selection in keystroke dynamics identity verification. In : *Neural Networks, 2003. Proceedings of the International Joint Conference on*. IEEE, 2003. p. 2253-2257.
- ZABALZA, Jaime, REN, Jinchang, REN, Jie, LIU, Zhe., and MARSHALL, Stephen. Structured covariance principal component analysis for real-time onsite feature extraction and dimensionality reduction in hyperspectral imaging. *Applied Optics*, 2014, vol. 53, no 20, p. 4440-4449.
- ZHANG, J., and FOODY, G. M. Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical and artificial neural network approaches. *International journal of remote sensing*, 2001, vol. 22, no 4, p. 615-628.
- ZHAO, Dongyu, ZHU, Shiping., and WANG, Fengchao. Lossy hyperspectral image compression based on intra-band prediction and inter-band fractal encoding. *Computers & Electrical Engineering*, 2016, vol. 54, p. 494-505.

- ZHONG, Ping., and WANG, Runsheng. Modeling and classifying hyperspectral imagery by CRFs with sparse higher order potentials. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, vol. 49, no 2, p. 688-705.
- ZHU, Guokang, HUANG, Yuancheng, LEI, Jingsheng, BI, Zhongqin, and XU, Feifei. Unsupervised hyperspectral band selection by dominant set extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, vol. 54, no 1, p. 227-239.
- ZHUO, Li, ZHENG, Jing, WANG, Fang, LI, Xia, AI, Bin., and QIAN, Junping. A genetic algorithm based wrapper feature selection method for classification of hyper spectral data using support vector maching. *Geographical Research*, 2008, vol. 27, no 3, p. 493-501.