

PANORAMA SUR LES METHODES DE CLASSIFICATION DES IMAGES SATELLITES ET TECHNIQUES D'AMELIORATION DE LA PRECISION DE LA CLASSIFICATION

O. El Kharki¹, J. Mechbouh², D. Ducrot³, M. Rouchdi⁴, j. M. Ngon⁵

1: Cellule Informatique, ENCG, Agadir, Maroc -- elkharki@hotmail.com

2: Craste-LF, Rabat, Maroc,

3: CESBIO, Toulouse, France.

4 : IAV Hassan II, Rabat, Maroc.

5 : Académie Internet, Université de Douala, Cameroun

Résumé

Au cours des dernières années, de nombreuses approches avancées de classification, tels que les réseaux de neurones artificiels, arbres de décision, les ensembles flous, etc. ont été largement appliquées à la classification des images satellites. Chaque méthode de classification a son propre mérite. Sélectionner une approche de classification appropriée pour une étude spécifique n'est pas facile. Différents résultats de classification peuvent être obtenus selon le(s) classificateur(s) choisi(s). Dans cet article, nous passons en revue diverses méthodes de classification avec une analyse et étude comparative. Nous présentons également les techniques pour améliorer la précision de la classification de la couverture terrestre.

Mots-clés : Télédétection, images satellites, classification, SVM, réseaux de neurones, classification floue, arbre de décision.

Abstract

In recent years, many advanced classification approaches, such as artificial neural networks, decision trees, fuzzy sets, have been widely applied for image classification. Each classification method has its own merits. Select an appropriate classification approach for a specific study is not easy. Different classification results can be obtained according to the selected classifier(s). In this paper, we review various methods of classification with an analysis and comparative study. We also present techniques to improve the accuracy of the classification of land cover.

Keywords: Remote sensing, satellite images, classification, SVM, neural networks, fuzzy classification, decision tree.

1. Introduction

La classification des données de télédétection a longtemps attiré l'attention de la communauté de télédétection, car les résultats de la classification sont des sources fondamentales pour de nombreuses applications environnementales et socio-économiques. Les scientifiques et les praticiens ont fait de grands efforts dans l'élaboration d'approches et de techniques de classification de pointe pour améliorer la précision de la classification (Stuckens et al., 2000 ; Pal et Mather, 2003 ; Gallego, 2004). Cependant, la classification des données de télédétection pour élaborer des cartes thématiques reste un défi en raison de nombreux facteurs, tels que la complexité du paysage de la zone d'étude, la sélection de données de télédétection, la sélection de méthodes de traitement d'images et des approches de classification. Ces facteurs affectent la précision de la classification. Bien que beaucoup de recherches antérieures soient particulièrement concernées par la classification d'images (Tso et Mather, 2009), une étude exhaustive des différentes approches de classification des images satellites n'était pas disponible à ce jour. L'émergence continue de nouveaux algorithmes et des techniques de classification ces dernières années nécessite un tel

examen, qui sera très utile pour guider ou sélectionner une procédure de classification appropriée pour une étude spécifique.

L'objectif de cet article est de présenter l'état de l'art des méthodes de classification avec une analyse et une étude comparative. Une attention particulière est accordée aux principales méthodes de classification avancées telles que les réseaux de neurones, les SVMs (Support Vector Machine), la classification floue et l'arbre de décision et aux techniques utilisées pour améliorer la précision de la classification.

2. Principe de la classification

La classification est un procédé algorithmique visant à diviser l'ensemble des éléments $\{p_i, i \in \{1, \dots, N\}\}$ d'une population P en un nombre C de partitions, selon un critère de similarité dépendant du contexte et du résultat désiré. En imagerie où P correspond à une image et p_i à un pixel, l'objectif est de produire des cartes thématiques mettant en valeur de manière simple et schématique différentes régions élémentaires d'une scène (figure 1). Une partition - ou classe - représente une famille de pixels ayant des propriétés communes et qui identifie ainsi une ou plusieurs zones présentant des caractéristiques physiques (type de sol,

de végétation, taux d'humidité) ou contextuelles propres. Cela rend la notion de similarité entre pixel double : spectrale et/ou spatiale. La qualité des résultats obtenus dépend grandement de la robustesse des classes, ces dernières doivent satisfaire aux trois exigences suivantes :

- Exhaustivité : tous les pixels de l'image doivent pouvoir être associés à une classe.
- Séparabilité : les classes doivent être suffisamment différenciables pour qu'un pixel ne puisse être associé qu'à une seule classe.
- Pertinence : l'ensemble des classes doit refléter l'information que l'utilisateur veut retirer de l'image.

Ces exigences doivent être vérifiées simultanément, les deux premières étant liées aux données tandis que la dernière dépend de l'utilisateur.

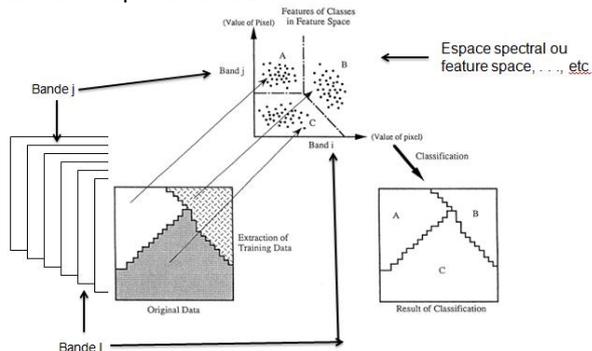


Figure 1 : Principe de la classification d'une image satellite.

La classification des images est un processus complexe qui peut être affecté par de nombreux facteurs. Une classification réussie nécessite de l'expérience et de l'expérimentation. L'analyste doit choisir une méthode de classification qui convient le mieux pour accomplir une tâche spécifique. En général, les méthodes de classification des images satellites peuvent être regroupées en supervisée et non supervisée, ou paramétrique et non paramétrique, ou dur et flou, ou par pixel (en anglais : per-pixel), sub-pixel (sous pixel) et orientée objet (object – oriented), figure 2.

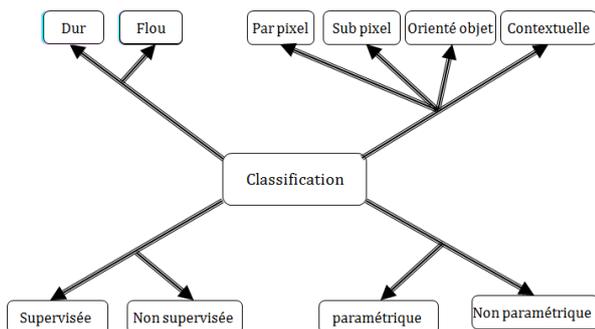


Figure 2 : Catégorisation des méthodes de classification des images satellites

3. Classification par pixel

L'approche pixel prend le pixel comme élément de référence. Le regroupement est réalisé selon le seul critère de ressemblance spectrale (Wong et al., 2003). Les algorithmes de classification par pixel peuvent être supervisés ou non supervisés ou les deux.

Dans le cas d'une classification non supervisée, l'image est partitionnée sans utiliser de connaissances a priori. Le classificateur crée des regroupements de pixels ayant des caractéristiques spectrales semblables. Il existe principalement deux grandes familles d'approches : la classification par partitionnement et la classification hiérarchique.

Bien que la sélection des données d'entraînement puisse être fastidieuse, une approche supervisée est préférée par la plupart des chercheurs, car elle donne généralement des définitions de classes plus précises par rapport aux approches non supervisées. Ce type de classification comprend en général trois étapes :

- 1- Etape d'entraînement au cours de laquelle des pixels de classes connues sont identifiés et caractérisés pour former les statistiques d'entraînement servant à décrire les classes.
- 2- Etape d'allocation des classes : les statistiques sont utilisées pour affecter chaque pixel de l'image à la classe avec laquelle il a la plus grande vraisemblance.
- 3- Etape de vérification : évalué la qualité de la classification.

Parmi les méthodes de classification conventionnelles supervisées, trois classificateurs statistiques sont d'usage général. Il s'agit de la méthode de parallélépipède, classificateur de la distance minimale, et l'algorithme du maximum de vraisemblance.

3.1 K-means

L'algorithme K-means (Linde et al., 1980) est le représentant le plus répandu de la famille par partitionnement. Cet algorithme est non supervisé. Il est à la fois rapide et simple mais outre le fait de devoir fixer au préalable le nombre de groupes, ses résultats présentent également une forte dépendance au choix des centroïdes initiaux, ce qui implique généralement de réitérer plusieurs fois le processus complet afin de maximiser les chances de converger vers un "meilleur" minimum local (Roussel, 2012).

3.2 Les algorithmes de classification hiérarchiques

Les algorithmes de classification ascendante ou descendante hiérarchique sont non supervisés. Ils génèrent, respectivement par fusion ou par division, une suite de partitions des observations allant de la partition mono-classe, où la même classe englobe la totalité des observations, à la partition où chaque observation correspond à une classe. L'objectif de ces outils est de proposer à l'analyste de choisir manuellement ou automatiquement la partition de classe la plus réaliste. Il en existe plusieurs variantes qui se distinguent essentiellement par leur critère de fusion ou de division des classes (distance euclidienne, angle spectral, etc.).

3.3 Parallélépipède

La méthode parallélépipède est supervisée. Elle est implémentée en définissant dans l'espace spectral des parallélépipèdes (i.e., l'hyper-rectangles : hyper-boîtes) pour chaque classe. Les limites du parallélépipède, pour chaque classe, peuvent être définies par les valeurs minimales et maximales des pixels dans une classe donnée, ou bien par un certain nombre d'écart-types autour des moyennes des classes. La règle de

décision doit simplement vérifier si un pixel dans l'espace spectral se trouve à l'intérieur de l'un des parallélépipèdes. La méthode parallélépipède est rapide et facile à mettre en œuvre, mais des erreurs peuvent survenir, particulièrement quand un pixel se trouve à l'intérieur de plus d'un parallélépipède ou à l'extérieur de tous les parallélépipèdes. Ces deux situations sont, en fait, susceptibles de se produire, car dans l'espace spectral, la distribution des vecteurs modèle est souvent très complexe. Par conséquent, il est difficile de fournir une performance de classification robuste utilisant cette simple méthode.

3.4 Distance minimale

Avec l'algorithme de distance minimale, la règle de décision pour l'affectation d'un pixel à une classe est la distance minimale entre la valeur du pixel et les centres de classes, mesurés par la distance euclidienne ou distance de Mahalanobis généralisée. Ces deux mesures de distance peuvent être décrites comme des coefficients de dissimilarité (i.e., la similitude entre deux objets i et j augmente lorsque la distance diminue). Pour chaque classe, le vecteur moyen spectral, ou centroïde de classe est déterminé à partir des données

d'entraînement. Si la distance de Mahalanobis est utilisée à la place de la distance euclidienne, on ajoute les paramètres de variance-covariance. Pour effectuer une classification par distance minimale, l'algorithme doit calculer la distance par rapport à chaque centroïde d'une classe, pour chaque pixel inconnu. Un pixel anonyme est affecté à la classe la plus proche, c'est-à-dire la plus proche de son centre de gravité.

La classification par la distance de Mahalanobis donne de bons résultats si l'image satellite présente des classes compactes et de forme convexe (ellipsoïde) (Lazar, 2008). Ce type de classificateur est mathématiquement simple et rapide en temps d'exécution, mais la base théorique peut ne pas être aussi robuste que celle du classificateur du maximum de vraisemblance. Une comparaison rapportée par Benediktsson et al. (1990) montre que la méthode de classification basée sur la distance de Mahalanobis est plus performante que celle basée sur la distance euclidienne. Mais sa précision de la classification reste faible par rapport à celle obtenue par l'algorithme du maximum de vraisemblance. Le tableau 1 présente une comparaison des méthodes de classification dirigée.

Classificateur	Avantages	Inconvénients
Parallélépipède	- Simple et rapide, - Aucun hypothèse sur la distribution statistique des classes.	- Les pixels sont classifiés à partir de la moyenne spectrale.
Distance minimale	- Pas de pixels non classés, - Classificateur rapide.	- Aucune considération à la covariance des classes.
Mahalanobis	- Tient compte de la covariance des classes.	- Sur-classification des signatures ayants des valeurs élevées dans la matrice de covariance. - Suppose à priori que la distribution des classes est normale.
Maximum de vraisemblance	- Relativement précise, - Tient compte de la variabilité des classes.	- Beaucoup de calcul, - Suppose à priori que la distribution des classes est normale, - Sur-classification des signatures ayants des valeurs élevées dans la matrice de covariance.

Tableau 1 : Comparaison des méthodes de classification dirigée

3.5 Maximum de vraisemblance

Le maximum de vraisemblance (MV) fait partie des classificateurs supervisés paramétriques. Il suppose que les données d'entraînement de chaque classe, dans chaque bande, ont une distribution particulière, généralement gaussienne et que les paramètres statistiques (par exemple vecteur moyen et la matrice de covariance) générés à partir des échantillons sont représentatifs. La classification par maximum de vraisemblance utilise des statistiques, comprenant le vecteur moyen et la matrice de covariance pour chaque classe. Pour placer un vecteur de mesures X d'un pixel inconnu dans une classe, la règle de décision du maximum de vraisemblance calcule pour chaque pixel la probabilité d'appartenance conditionnelle à chaque classe. Puis, elle affecte le pixel à la classe qui offre la probabilité la plus élevée.

Le MV peut être le classificateur paramétrique le plus couramment utilisé dans la pratique, en raison de sa robustesse et de sa disponibilité dans presque tous les logiciels de traitement d'images de télédétection. Toutefois, l'hypothèse de la distribution normale est

souvent violée, en particulier dans les paysages complexes. En outre, des échantillons insuffisants, non-représentatifs, ou qui ont une distribution multimodales peuvent encore introduire l'incertitude sur la procédure de classification. Avec des classificateurs non paramétriques, l'hypothèse d'une distribution normale de l'ensemble de données n'est pas nécessaire. Les recherches ont montré que beaucoup de classificateurs non paramétriques peuvent donner de meilleurs résultats de classification que les classificateurs paramétriques dans des paysages complexes (Foody, 2002). Parmi les approches de classification non paramétriques les plus couramment utilisées, nous pouvons citer : les réseaux de neurones artificiels, arbres de décision et les machines à vecteurs de support (SVM).

3.6 Réseaux de neurones

Les réseaux de neurones sont constitués d'un ensemble de neurones artificiels ou nœuds qui sont analogues aux neurones biologiques. Ils sont issus d'une tentative de conception d'un modèle

mathématique très simplifié du cerveau humain en se basant sur notre façon d'apprendre et de corriger nos erreurs.

Dans un réseau de neurones multicouches, les neurones sont organisés en couches successives (Figure 3).

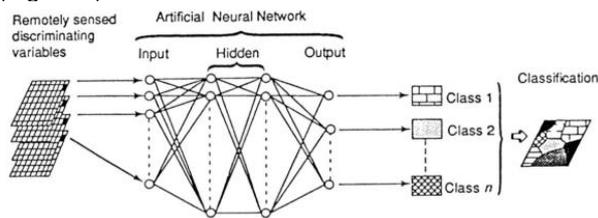


Figure 3 : Architecture d'un réseau de neurones multicouche en télédétection

Chaque neurone d'une couche reçoit des signaux de la couche antérieure et transmet le résultat aux neurones de la couche postérieure en suivant un seul sens de propagation de l'information (entrée - sortie). Avec cette configuration, la couche d'entrée reçoit les variables d'entrée et la couche de sortie fournit les résultats. Les autres couches qui se trouvent entre la couche d'entrée et la couche de sortie sont appelées couches cachées. L'approche des réseaux de neurones artificiels (ANN : Artificial Neural Networks) a été largement adoptée au cours des dernières années. L'utilisation de cette approche comme outil de traitement de données de télédétection est principalement motivée par les raisons suivantes :

- 1) la capacité d'opérer avec une plus grande précision que les autres techniques, à l'exemple des classificateurs statistiques, particulièrement lorsque l'espace des caractéristiques est complexe, et que les sources de données ont des distributions statistiques différentes.
- 2) la possibilité de faire le parallélisme, ce qui permet le traitement en temps réel de très grands ensembles de données.
- 3) La résistance aux pannes (si un neurone ne fonctionne plus, le réseau ne se perturbe pas).
- 4) Temps de classification rapide.

Au regard de ces nombreux atouts, il est clair que l'une des principales opportunités offertes par les réseaux de neurones est la possibilité de traiter efficacement les grandes quantités de données de télédétection qui sont actuellement disponibles. Toutefois, l'utilisateur doit déterminer l'architecture du réseau, et aussi définir des paramètres tels que la vitesse d'apprentissage, qui affectent le temps d'entraînement, les performances et la vitesse de convergence d'un réseau neuronal. Il n'y a pas de règles claires pour aider à la conception du réseau, et que des règles empiriques (ou heuristiques) existent pour guider les utilisateurs dans le choix des paramètres du réseau.

Les classificateurs neuronaux s'appliquent aux deux types de classifications : supervisée et non supervisée (Sonka et al., 1998). Le modèle de réseaux de neurones le plus communément utilisé pour la classification des images de télédétection est le perceptron multicouche. Ces réseaux sont constitués d'un minimum de trois couches de neurones. Afin de paramétrer les poids associés à ces derniers, l'algorithme d'apprentissage le plus répandu est la rétro-propagation du gradient (Rumelhart, 1986). Ce dernier propage l'erreur quadratique obtenu sur les échantillons d'apprentissage et détermine la contribution des poids à l'erreur générale, puis il les

corrige de manière à se rapprocher le plus possible du résultat souhaité. Dans un cadre plus général, l'apprentissage consiste en un entraînement du réseau. On présente au réseau des entrées et on lui demande de modifier sa pondération de telle sorte que l'on retrouve la sortie correspondante. Un perceptron multicouche, combiné à l'analyse texturale, a été utilisé par Akono et al. (1996) pour la classification de la forêt des Mangroves dans la côte littorale camerounaise. Ndi et al. (1997) ont utilisé un perceptron multicouche, combiné à l'algorithme de classification floue fuzzy c-means pour effectuer la classification de la forêt des mangroves dans la même région. Ces travaux ont produit des classifications avec une précision de plus de 95%.

3.7 Support Vecteurs Machines (SVM)

Les machines à vecteurs de support, ou séparateurs à vaste marge (Support Vector Machine : SVM) sont des méthodes de classification par apprentissage supervisée, elles ont été introduite depuis plusieurs années en théorie d'apprentissage pour la classification et la régression (Vapnik, 1998 ; Theodoridis et al., 2010). Plus récemment, cette méthode est appliquée à la classification des images de télédétection (Melgani et Bruzzone, 2004; Roli et Fumera, 2001). Cependant, les SVM sont des classificateurs qui permettent de traiter des problèmes de discrimination non-linéaire, et de reformuler un problème de classement en un problème d'optimisation quadratique. Cette optimisation est relative à la notion de « marge maximale ». La marge est la distance entre la frontière de séparation et les échantillons les plus proches. En effet, c'est cette frontière de séparation (ou hyperplan) qui doit être optimale, et pour cela, un bon choix de l'ensemble d'apprentissage s'impose.

Afin de pouvoir traiter les cas de données qui ne sont pas linéairement séparables, les SVM permettent de transformer l'espace de représentation des données d'entrée (espace spectral) en un espace de plus grande dimension dans lequel il est probable qu'il existe un séparateur linéaire (i.e., droite, plan, hyperplan). Ceci est réalisé grâce à une fonction noyau (kernel). Son avantage réside en ce qu'elle donne de bons résultats même lorsque le nombre d'échantillons utilisés pour l'apprentissage est petit. Le seul défaut réside dans la difficulté de choisir puis paramétrer un noyau adapté à un tel jeu de données de télédétection.

Comme les classificateurs ANN, les paramètres définis par l'utilisateur pour SVM ont une forte influence sur le résultat de la classification. Cela peut restreindre les capacités de SVM pour résoudre les problèmes de classification plus complexes. Cependant, les performances de haut niveau de SVM ont attiré l'attention et l'intérêt des chercheurs. Giorgos (2011) a présenté un état de l'art sur les SVM en télédétection. Des études récentes ont montré que l'utilisation de SVM pour traiter des problèmes de classification des images satellites peut donner une plus grande précision que les autres classificateurs et n'exigent que moins d'échantillons. Par exemple, Foody et Mathur (2004) ont montré que seulement un quart des échantillons d'entraînement pour une image SPOT HRV était suffisant pour produire la même précision qu'un classificateur à deux cultures. D'autres études ont suggéré que SVM peut donner de bons résultats pour la classification des images hyperspectrales et des

résultats supérieurs ont été rapportés par rapport aux algorithmes traditionnels de classification des données de télédétection tels que MV, k-plus proche voisin, et ANN (Huang et al., 2002; Melgani et Bruzzone, 2004; Pal et Mather, 2005). La propriété la plus attrayante de SVM est la grande capacité de généralisation avec un nombre relativement restreint des échantillons d'entraînement (Bishop, 2006; Pal et Mather, 2005). Tous ces avantages du SVM ont fait de cette méthode, le classificateur le plus employé pour des données hyperspectrales dans la dernière décennie (Licciardi et al., 2009).

3.8 Arbres de décisions

Les arbres de décisions sont des méthodes de classification par apprentissage supervisée et non supervisée. La conception d'un arbre de décision (AD) est fondée sur la connaissance des propriétés spectrales de chaque classe et des relations entre les classes. Le principal avantage d'utiliser une structure arborescente pour effectuer des décisions de classification est que cette structure peut être considérée comme une zone blanche, qui par rapport à la méthode ANN, est plus facile à interpréter et à comprendre la relation entre les entrées et les sorties.

Un arbre de décision est composé d'une racine qui est le point de départ de l'arbre, des nœuds, des branches qui relient: la racine avec les nœuds, les nœuds entre eux et les nœuds avec les feuilles. En télédétection, les nœuds non terminaux représentent les attributs avec les seuils de séparabilité, les classes de pixels obtenues après traitement jouent le rôle des feuilles ou nœuds terminaux (figure 4).

Le procédé de classification est mise en œuvre par un ensemble de règles qui déterminent la voie à suivre, en partant du nœud racine et se terminant à un nœud terminal, ce qui représente l'étiquette du pixel à classer. A chaque nœud non terminal, une décision doit être prise sur le chemin vers le prochain nœud. Il est évident que la nature des décisions qui sont fixés et la séquence d'attributs se produisant à l'intérieur d'un arbre aura une incidence sur les résultats de la classification. Ainsi, l'efficacité et la performance de cette approche est fortement affectée par l'algorithme introduisant l'arbre de décision.

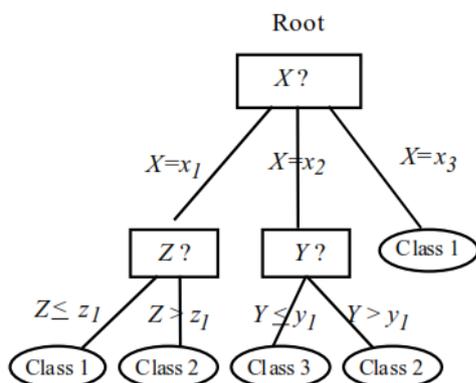


Figure 4 : Principe de la classification par arbre de décision

Deux approches de la conception d'un arbre de décision peuvent être envisagées. L'une est basée sur la connaissance de l'utilisateur et repose uniquement sur l'interaction de l'utilisateur. C'est ce qu'on appelle l'approche manuelle de la conception. Dans la

procédure de conception manuelle, les statistiques pour toutes les classes sont d'abord calculées, et un graphique de la gamme spectrale dans chaque bande est construit. A partir du graphique, et des paramètres statistiques, des estimations des limites de décision sont dérivés et l'arbre est conçu manuellement dans des classes séparées de façon hiérarchique. La construction d'un arbre, par des méthodes de conception manuelle prend du temps et peut ne pas fournir des résultats satisfaisants, en particulier lorsque le nombre de classes est grand et il y a un chevauchement spectral entre les classes. La seconde approche utilise une procédure automatique. L'intérêt d'utiliser des méthodes automatiques pour la conception et l'utilisation de classificateurs d'arbres de décision a rapidement augmenté ces dernières années et les performances relatives sont prometteuses. Ces classificateurs cherchent en premier lieu la bande la plus hétérogène parmi les bandes existantes en calculant les indices d'hétérogénéité : entropie pour l'algorithme ID.3 (itératif Dichotomizer 3), entropie et ratio pour l'algorithme C4.5 et indice de Gini pour l'algorithme CART (Classification and Regression Tree), Cette bande est ensuite introduite dans la racine. La classification dans la bande utilise également les seuils de séparabilité (les valeurs maximales ou minimales de chaque classe selon les cas).

Swain et Hauska (1977) ont proposé une technique de recherche heuristique basée sur une fonction d'évaluation mathématique pour résoudre des problèmes qui sont plus complexes. D'autres stratégies pour la conception optimale d'arbre sont décrites par Kulkarni et Laveen (1976) et Kurzynski (1983). La conception d'un classificateur à arbre optimum, dépend du choix de la structure d'arbre, des attributs utilisés dans chaque nœud terminal, et des règles de décision pour effectuer la classification au niveau de chaque nœud non terminal (Tso et Mather, 2009). Un examen approfondi des différentes méthodes utilisées pour élaborer les classificateurs d'arbre de décision est donné par Safavian et Landgrebe (1991).

L'utilisation des techniques d'arbre de décision pour classer les images de télédétection a connu une croissance rapide ces dernières années. Hansen et al. (1996) ont appliqué l'arbre de décision dans la classification de la couverture terrestre, et d'autres auteurs ont suivi leur exemple (Muchoney et al., 2000; Pal et Mather, 2003 ; Kandrika et Roy, 2008, Wen et al., 2008). Un certain nombre d'auteurs ont procédé à la comparaison d'arbre de décision avec d'autres classificateurs. Gahegan and West (1998) ont trouvé que le classificateur MV n'est pas aussi performant qu'un arbre de décision dans l'étiquetage des pixels d'une image d'une zone de test complexe. Muchoney et al. (2000) comparent un réseau de neurones, le classificateur MV, et un arbre de décision dans une étude de la végétation et la cartographie de la couverture terrestre en Amérique centrale. Ils ont conclu que l'arbre de décision produit des résultats supérieurs. Otukei et Blaschke (2010) ont comparé AD, MV et SVM pour l'évaluation des changements de la couverture terrestre à l'aide de Landsat TM et ETM+. Ils ont trouvé que la méthode AD est plus performante par rapport aux autres. Le tableau 2 présente une comparaison entre les trois méthodes de classification non paramétrique.

Plusieurs études sont concentrées sur l'amélioration de la précision des classificateurs AD en utilisant des algorithmes avancés et l'intégration des données

auxiliaires et des données multisources. Evrendilek et Gulbeyaz (2011) ont utilisé AD avec le climat et les

données topographiques pour la cartographie à partir des images satellites MODIS.

Classificateur	Avantages	Inconvénients	Supervisée et/ou non
ANN	<ul style="list-style-type: none"> - Parallélisme, - Résistance aux pannes, - Rapide au niveau classification. 	<ul style="list-style-type: none"> - Apprentissage lent, - Les ANN sont sémantiquement pauvres, - Problème de convergence stochastique - Le résultat de la classification dépend de l'architecture du réseau. 	les deux
SVM	<ul style="list-style-type: none"> - Bases théoriques très solides, - N'exige que peu échantillons, - Nécessite moins d'effort pour designer l'architecture adéquate. 	<ul style="list-style-type: none"> - Les paramètres définis par l'utilisateur ont une forte influence sur le résultat de la classification. 	oui
AD	<ul style="list-style-type: none"> - Facile à interpréter et à comprendre la relation entre les entrées et les sorties. - Bonne efficacité de calcul. 	<ul style="list-style-type: none"> - La précision de la classification dépend entièrement de la conception de l'arbre de décision et des attributs utilisés. 	les deux

Tableau 2 : Comparaison des méthodes de classification paramétrique

4. Approches sous-pixels (Subpixel)

La plupart des méthodes de classification sont basées sur l'approche pixel (l'approche dur), dans laquelle chaque pixel est classé dans une seule classe et les catégories de couverture terrestre sont mutuellement exclusives. En raison de l'hétérogénéité des paysages et la limitation de la résolution spatiale de l'imagerie satellitaire, les pixels mixtes sont fréquents dans les données basses et moyennes résolutions spatiales. La présence de pixels mixtes a été reconnue comme un problème majeur, affectant l'utilisation efficace des données de télédétection dans les classifications par pixel (Fisher, 1997 ; Cracknell, 1998). Les approches de classification sub-pixels (sous-pixel ou démelange) ont été développées pour fournir une représentation plus appropriée dans un environnement hétérogène. Une représentation floue, dans laquelle chaque emplacement est composé des appartenances multiples et partielles de toutes les classes candidats, est nécessaire. La technique des ensembles floue (Maselli et al., 1996 ; Zhang et Kirby, 1999 ; Zhang et Foody, 2001 ; Shalan et al., 2003 ; Smara et al., 2003) et la classification par l'analyse de mélange spectral (SMA) (Roberts et al., 1998 ; Rashed et al., 2001 ; Lu et al., 2003) sont les méthodes les plus utilisés pour résoudre le problème de pixel mixte. Un inconvénient majeur de la classification par l'approche subpixel réside dans la difficulté d'évaluer la précision.

SMA a longtemps été reconnu comme une méthode efficace pour traiter le problème de pixel mixte. Elle évalue chaque spectre de pixel comme une combinaison linéaire d'un ensemble de composantes spectrales (Adams et al., 1995). La sortie du SMA est généralement présentée sous la forme d'images de fractions, avec une image pour chaque spectre de l'endmember, représentant les proportions des composantes spectrales au sein du pixel de la zone. Des recherches antérieures ont démontré que la méthode SMA est utile pour améliorer la précision de la classification (Lu et al., 2003) et est particulièrement important pour améliorer l'estimation des superficies des classes d'occupation du sol basée sur des données de grande résolution spatiale.

La classification floue (douce : élastique : soft) qui applique le concept d'ensembles flous assigne, non pas à un pixel une étiquette relative à une classe unique, mais son degré d'appartenance à chacune des classes. De cette façon, un même pixel peut appartenir à plusieurs classes. La théorie des ensembles flous (Zadeh, 1965), fournit un cadre conceptuel pour la résolution des problèmes de classification dans un environnement ambigu. C'est pour cela que plusieurs recherches ont contribué à utiliser cette théorie sur les algorithmes classiques de classification, supervisée ou non supervisée.

La plupart des algorithmes de classification floue analytique sont dérivées de l'algorithme Floue C-Means (FCM) de Bezdek (1981). L'algorithme FCM et ses dérivés ont été mis en œuvre avec succès dans de nombreuses applications, telles que la classification des formes et la segmentation d'images. L'algorithme FCM consiste à trouver l'appartenance exacte d'un pixel donné à une classe. Au départ, à chaque pixel est assignée une valeur qui représente son degré d'appartenance à chaque classe. Ce degré varie entre 0 et 1 (Wang, 1990): c'est la fuzzyfication. On applique la règle floue choisie, cette règle gère la défuzzyfication du système en affectant chaque pixel à une seule classe à savoir celle dont il a le plus grand degré d'appartenance. Le principe consiste à trouver une partition de l'image caractérisée par une matrice de partition qui caractérise le degré d'appartenance d'un pixel à une classe. En générale, les performances des méthodes de classification floue est supérieure à celle des versions dur correspondantes et ils sont moins susceptibles de s'en tenir à un minimum local (Bezdek, 1981).

5. Approche orientée objet

Les images de haute résolution spatiale ont été de plus en plus utilisées pour la classification des couvertures terrestre, mais la variation spectrale élevée au sein de la même classe, la confusion spectrale entre les différentes couvertures terrestre, et le problème de l'ombre rendant les classificateurs par pixel de moins en moins performants (Blaschke et al., 2001). L'approche de classification orientée objet (COO) est

conçu pour traiter le problème de l'hétérogénéité de l'environnement, elle ne traite plus le pixel de manière isolée mais des groupes de pixels (objets) dans leur contexte. Deux étapes sont impliquées dans une COO: la segmentation d'image et la classification (El kharki et al., 2012a ; Jensen, 2004). La segmentation des images fusionne les pixels en objets. Puis des données spectrales, spatiales et contextuelles sur ces objets sont recueillies afin d'orienter la classification. Dans le processus de création d'objets, une échelle détermine l'apparition ou l'absence d'une classe d'objet, et la taille d'un objet ce qui affecte le résultat de la classification (El kharki et al., 2014b).

Cette approche s'est avérée être en mesure de fournir de meilleurs résultats de classification que les approches de classification par pixel, en particulier pour les données de haute résolution spatiale (Thomas et al., 2003; Benz et al., 2004; Laliberte et al., 2004; Yu et al., 2006; Stow et al., 2007; Mallinis et al., 2008).

6. Approches contextuelles

En plus de l'orientée objet, les classificateurs contextuels ont été également mis au point pour faire face au problème des variations spectrales intra-classe (Gong et Howarth, 1992 ; Kartikeyan et al., 1994 ; Flygare, 1997 ; Sharma et Sarkar, 1998 ; Keuchel et al., 2003; Magnussen et al., 2004). La classification contextuelle exploite les informations spatiales entre pixels voisins pour améliorer les résultats de la classification (Flygare, 1997 ; Stuckens et al., 2000 ; Hubert-Moy et al., 2001 ; Magnussen et al., 2004).

Une première approche consiste à calculer diverses caractéristiques spectrales et spatiales pour chaque pixel de l'image de télédétection puis à procéder à une classification par pixel. Les deux types d'informations peuvent être classés indépendamment, résultant en une carte de classification spatiale et une carte de classification spectrale qui sont fusionnées selon diverses règles. C'est ainsi que procèdent Benediktsson et al. (2004), utilisant un réseau de neurones pour classer un ensemble de profils morphologiques préalablement calculés sur les premières composantes principales de l'image et une approche type maximum de vraisemblance pour classer des bandes spectrales issues d'une réduction de dimension par NWF (Nonparametric Weighted Feature Extraction). Les caractéristiques peuvent être également concaténées les unes dans les autres pour former un unique jeu de données, auquel on applique une classification vectorielle. Lepisto (2003) adopte cette approche avec des caractéristiques spatiales de type bancs de filtres ou coefficients de Haralick et un classificateur de type k-plus proches voisins. Les bancs de filtres sont des caractéristiques spatiales obtenues par filtrage directionnel de l'image au moyen de divers outils de traitement du signal tels que les filtres de Gabor (Dunn et al., 1994) ou les transformées en ondelettes (Kuo et al., 1993). Les coefficients de Haralick sont quant à eux obtenus en étudiant les transitions de niveaux de gris pour un décalage de pixels donné. Palmason et al. (2005) et Fauvel et al. (2008) font de même avec des profils morphologiques, un troisième type de caractéristique spatiale obtenue par filtrage morphologique d'une image hyperspectrale. Les classificateurs utilisés sont alors respectivement un réseau de neurones et une machine à vecteurs supports. Camps-Valls et al. (2005) ont proposé les

noyaux composés pour la classification d'images hyperspectrales (Camps-Valls et al., 2006). Dans un premier temps, pour chaque pixel, des vecteurs de caractéristiques spectrales et spatiales sont extraits séparément. Par exemple, la moyenne et l'écart type (par bande spectrale) peuvent être estimés sur le voisinage le plus proche du pixel et être employés comme des caractéristiques spatiales. Les vecteurs de caractéristiques spectrales et spatiales sont utilisés pour calculer deux matrices de noyau différentes. Ensuite, les deux matrices de noyau sont combinées en employant une des techniques proposées des noyaux composés, telle que la concaténation des vecteurs de caractéristiques, un noyau d'addition directe (direct summation kernel), un noyau d'addition pondérée (weighted summation kernel), un noyau "cross-information" et des noyaux pour la polyvalence améliorée. Cette méthodologie a montré de bonnes performances pour l'intégration de l'information contextuelle dans la classification et a été utilisée en considérant des systèmes de voisinage plus avancés.

La seconde approche c'est la modélisation du contexte spatial par les champs aléatoires de Markov (MRFs : Markov Random Fields). Les MRFs constituent une approche probabiliste classique permettant de prendre en compte les interactions spatiales entre les pixels d'une image. Ils sont appliqués couramment pour intégrer le contexte spatial dans des problèmes de classification d'image (Dubes, 1993). Les méthodes ICM (Iterative Conditional Modes) sont les plus fréquemment utilisés dans la classification contextuelle (Magnussen et al., 2004), et ont prouvé leur efficacité dans l'amélioration des résultats de classification. Pony et al. (2000) et Jackson et Landgrebe (2002) ont proposé des classificateurs statistiques itératifs spectro-spatiaux pour des données hyperspectrales, qui utilisent une modélisation par MRFs. En conclusion, les méthodes utilisant MRFs se sont avérées être un outil puissant pour l'analyse contextuelle d'image. Cependant, ils comprennent une étape itérative d'optimisation, qui est longue en temps de calcul.

Enfin, une dernière approche consiste à tirer parti de l'information de forme contenue dans les images par l'intermédiaire d'une carte de segmentation réalisée préalablement. Dans (Tarabalka, 2009 ; 2010) le résultat d'une classification par SVM est régularisé au moyen d'une carte de segmentation et d'un vote majoritaire.

7. Combinaison de classificateurs

De nombreux chercheurs ont publié des comparaisons entre les différentes méthodes de classification des images de télédétection. Il semble que la plupart des méthodes sont complémentaires. Par conséquent, un certain intérêt a été montré à combiner les résultats obtenus en utilisant différentes règles de décision afin d'améliorer la classification globale. Différentes approches de combinaison de classificateurs ont été conçues et il a été démontré que l'intégration de deux ou plusieurs classificateurs permet d'améliorer la précision de la classification par rapport à l'utilisation d'un seul classificateur (Benediktsson et Kanellopoulos, 1999 ; Steele, 2000 ; Huang et Lees, 2004).

De nombreux chercheurs ont étudié le schéma de la combinaison de plusieurs classificateurs pour générer un résultat unique. Les études théoriques et empiriques

indiquent qu'une bonne combinaison est celle où les classificateurs sont précis et en même temps sont en désaccord sur certaines parties de l'espace d'entrée (Ceamanos et al., 2010). Deux approches populaires ont été utilisées dans la combinaison des résultats des différents classificateurs qui sont Bagging et Boosting. Bagging repose sur la méthode bootstrap pour améliorer la prédiction d'un classificateur (Breiman et al., 1996). Boosting est une méthode de production d'une règle de prédiction très précise (Freund et al., 1996). Certaines études antérieures ont démontré que le vote à la majorité de la prédiction de classificateur est une stratégie efficace. Dans ce schéma, un pixel est classé comme la classe qui a été choisi par la majorité des classificateurs. Foody et al. (2007) ont utilisé la règle vote majoritaire pour intégrer plusieurs classificateurs binaires pour la cartographie d'une classe spécifique. Doan et Foody (2007) ont étudié la combinaison de méthodes de classification floues des images de télédétection et ont trouvé que la combinaison des classifications pourrait améliorer la précision. Fauvel et al. (2006) ont fusionné les résultats de la classification issus d'un réseau de neurones et un classificateur flou en utilisant la théorie des ensembles flous, et la méthode proposée améliore la précision de la classification par rapport à l'utilisation séparée des différents classificateurs. Waske et al. (2007) ont fusionné SVM pour la classification des images radars de type SAR et des images multispectrales de zones agricoles, les résultats de l'approche proposée sont supérieures à toutes les autres techniques, (SVM standard).

8. Aperçu des principales techniques pour l'amélioration de la performance de classification

Différentes données de télédétection ont des variations spatiales, spectrales, radiométriques et temporelles, ainsi que des différences dans la polarisation. L'exploitation de ces caractéristiques est un moyen efficace pour améliorer la précision de la classification (Lu et Weng, 2005 ; Tso et Olsen, 2005). D'une manière générale, la réponse spectrale est l'information la plus utilisée pour la classification de la couverture terrestre. La texture et l'information contextuelle deviennent aussi de plus en plus importante grâce à la disponibilité des données de haute résolution spatiale (Lu et al., 2008). Ce chapitre aborde les principales techniques utilisées pour améliorer les performances de la classification de la couverture terrestre.

8.1 Utilisation de l'information spatiale

La résolution spatiale d'une image détermine le niveau de détail que l'on peut observer sur la surface de la Terre. L'information spatiale joue un rôle important dans l'amélioration de la précision de la classification des couvertures terrestres, en particulier lorsque les images à haute résolution spatiale sont utilisées (Goetz et al., 2003; Stow et al., 2007; Zhou et al., 2008). La texture et le contexte sont les deux aspects de mesure de la structure spatiale d'une image. Ils sont fréquemment utilisés dans l'interprétation d'une image et sont généralement extraits directement à partir de cette dernière.

L'analyse de texture apporte une information sur l'agencement spatial des niveaux de gris, qui est

nécessaire pour pallier aux faiblesses des méthodes de classification par pixel, surtout pour les images de haute résolution spatiale. Les principales approches de texture peuvent être classées en quatre groupes en fonction de leurs différentes origines théoriques (Tso, 2009). La première approche définit les caractéristiques de texture qui sont dérivées à partir du spectre de puissance de Fourier de l'image, par l'intermédiaire d'un filtrage dans le domaine fréquentiel. La deuxième approche est basée sur des statistiques qui mesurent les propriétés locales qui sont censés être liés à la texture, par exemple, la moyenne locale ou l'écart type. La troisième approche est l'utilisation de la densité de probabilité de niveau de gris commun (Haralick, 1973). L'approche finale est basée sur la modélisation de l'image en utilisant des hypothèses, telles que (1) l'image en cours de traitement possède des propriétés fractales (Mandelbrot, 1977, 1982), ou (2), il peut être modélisée en utilisant un modèle de champ aléatoire (Frankot et Chellapa, 1987).

L'utilisation de la texture, en plus des caractéristiques spectrales dans la classification d'images pourrait s'attendre à entraîner un certain niveau d'amélioration de la précision, en fonction de la résolution spatiale de l'image et la taille de la zone homogène à classer (Coburn et Roberts, 2004; Ouma et al., 2006).

Dans l'interprétation de l'information visuelle, le contexte est très important. Il peut être dérivé d'attributs spectraux, spatiaux ou même temporels. L'utilisation appropriée de contexte permet l'élimination des ambiguïtés, la récupération des informations manquantes, et la correction des erreurs (Magnussen et al., 2004). En utilisant le contexte, les pixels ne sont pas traités de façon isolée mais sont considérés comme ayant une relation avec leurs voisins. Ainsi, la relation entre le pixel auquel on s'intéresse et ses voisins est considérée comme étant statistiquement dépendante. Une façon simple d'utiliser le contexte est sous la forme d'une fenêtre du filtre de la majorité (Gurney, 1980, 1981). Bien qu'il y ait une certaine amélioration, l'augmentation de la précision de la classification n'est pas impressionnante. Des modèles statistiques plus complexes de contexte spatial sont discutés par Haslett (1985), Solaiman et al. (1996), et Nishii et Eguchi (2005).

Les Champs aléatoires de Markov sont un outil utile pour caractériser l'information contextuelle, ils ont été largement utilisés dans la segmentation et la restauration d'image (Magnussen et al., 2004; Tso et Olsen, 2005).

8.2 Intégration des données des capteurs différents

La fusion ou l'intégration de données multicapteurs ou multirésolutions profite des atouts de données d'images distinctes pour améliorer l'interprétation visuelle et / ou d'améliorer les performances de l'analyse quantitative (Ducrot, 2005 ; Klonus et Ehlers, 2007; Lu et Weng, 2007). En générale, la fusion de données de télédétection implique deux grandes procédures: (1) registration géométrique de deux ensembles de données, et (2) mélange de contenu de l'information spectrale et spatiale pour générer un nouvel ensemble de données qui contient l'information amélioré à partir des deux ensembles originaux (Lu et Weng, 2007). La registration précise entre deux ensembles de données est extrêmement importante pour extraire précisément le contenu de l'information à partir de deux ensembles de données, en particulier pour les éléments linéaires

comme les routes et les rivières. La fusion de données peut incorporer des données de différentes résolutions spatiales, telles que Quickbird, IKONOS, SPOT et Landsat ETM (Enhanced Thematic Mapper Plus), y compris les données multispectrales et panchromatiques (Ehlers et al., 2010; Zhang, 2010). Les données de différents capteurs peuvent également être combinés, comme Landsat TM et radar ou hyperspectrales et lidar (Light Detection and Ranging) (Lucas et al., 2006; Ali et al., 2009).

En générale, les techniques de fusion de données de télédétection peuvent être classées en trois niveaux: niveau pixel, niveau attribut et niveau décision (Chitroub, 2010). La fusion au niveau pixel est la combinaison de données brutes provenant de sources multiples en données de résolution unique. La fusion au niveau attribut extrait différents éléments tels que les contours, les lignes, les paramètres de texture, etc., provenant de différentes sources de données, puis les combine en une ou plusieurs cartes d'attributs qui peuvent être utilisés à la place des données d'origine pour un traitement ultérieur. Ceci est particulièrement important lorsque le nombre de bandes spectrales disponibles devient grand de telle sorte qu'il est impossible d'analyser chaque bande séparément. La fusion au niveau décision est basée sur la combinaison des résultats de plusieurs algorithmes pour obtenir une décision finale fusionnée. Zhang (2010) a souligné les techniques de fusion de données multisources et a discuté leurs tendances.

8.3 L'utilisation de données multitemporelles

La résolution temporelle fait référence à l'intervalle de temps dans lequel un satellite revisite la même zone géographique. L'utilisation de données de télédétection recueillies au cours des différentes saisons s'est avéré utile dans l'amélioration de la précision de la classification, en particulier pour la classification des cultures et de la végétation (El kharki et al., 2012b ; Wolter et al., 1995 ; Ducrot, 2005 ; Lunetta et Balogh, 1999 ; Oetter et al., 2000 ; Guerschman et al., 2003). Par exemple, Lunetta et Balogh (1999) ont comparé des images Landsat 5 TM, monodate et deux dates pour la cartographie des zones humides dans le Maryland et le Delaware. Ils ont constaté que les images multitemporelles fournies de meilleurs taux de classification par rapport à l'image monodate. Une précision globale de classification de 88 % a été atteinte à partir d'images multitemporelles, comparativement à 69 % pour l'image monodate.

8.4 Utilisation des données auxiliaires

Les données auxiliaires, telles que la topographie, les sols, les routes et les données de recensement, peuvent être combinées avec des données de télédétection pour améliorer les performances de la classification (El kharki et al., 2002). Les données topographiques (MNT) sont révélés être précieux dans l'amélioration de la précision de la classification de la couverture terrestre dans les régions montagneuse (Franklin et al., 1994). Dans les études urbaines, les données relatives aux systèmes humains tels que la distribution de la population et la densité des routes sont souvent incorporés dans les classifications urbaines (Epstein et al., 2002; Lu et Weng 2006).

Les différentes approches proposées pour la combinaison de l'information spectrale et auxiliaires comprennent: l'approche logique, la stratification, l'utilisation de la probabilité a priori et la classification orientée objet.

L'approche logique consiste à ajouter les données auxiliaires en tant que caractéristiques supplémentaires, de telle sorte que le vecteur pixel est étendu par l'ajout de ces informations externes (Gercek, 2004).

L'approche stratification implique la subdivision ou la segmentation de la zone d'étude en zones plus petites ou strates, sur la base de règles issues de la connaissance externe, de sorte que chaque strate peut être traitée indépendamment (De Bruin et Gorte, 2000). Ce processus est exécuté avant la classification, dans le but d'augmenter l'homogénéité de l'ensemble de données à classer, ou pour séparer des objets qui sont spectralement similaires.

Dans la plupart des procédés de classification, la probabilité d'appartenance à une classe d'un pixel donné est supposé égale pour toutes les classes. Toutefois, si l'information sur une zone montre la préférence de certaines classes pour des emplacements particuliers dans le terrain, cette information peut être exprimée en termes de probabilités a priori d'occurrence pour chaque classe et cette information peut être incorporé dans le processus de classification.

La dernière approche est l'utilisation des informations de limites afin de définir des objets avant leur classification. Cette stratégie de classification est adaptée pour des zones éparées (par exemple, les zones constituées de terres agricoles). Dans les applications de cultures agricoles, les objets sont des champs, dont les limites peuvent être dérivées soit par numérisation des cartes papier de l'échelle appropriée, ou en appliquant un algorithme de détection des bords de l'image. Chaque objet (champ) est caractérisé par des paramètres statistiques globaux, et est représenté par un vecteur unique dans l'espace des caractéristiques (Tso and Olsen, 2005).

9. Synthèse sur les méthodes de classifications des images satellites

Les approches de classification des images satellites peuvent être regroupés en approche par pixel, sub-pixel, orienté objet, contextuelle et à base de connaissances ou systèmes experts (tableau 3). La classification par pixel est encore la plus utilisée dans la pratique. Cependant, la précision de cette approche ne peut pas répondre à l'exigence des utilisateurs en raison de l'impact du problème du pixel mixte présent dans les images satellites basse et moyenne résolution spatiale. Les algorithmes de sub-pixel ont le potentiel de traiter le problème de pixel mixte, et peuvent atteindre une précision élevée pour des images moyenne et basse résolution spatiale. Pour les images haute et très haute résolution spatiale, le problème de pixel mixte est réduit. Cependant, la présence de l'ombre et la variation spectrale intra-classes rendent l'approche sub-pixel inadapté pour la classification de ces données. Les approches de classification orientée objet et contextuelles sont les plus appropriés pour les données de haute et très haute résolution spatiale. Lors de l'utilisation des données multisources (i.e. la combinaison de signatures spectrales, la texture et le

contexte et les données auxiliaires) les classificateurs non paramétriques avancés, tels que les réseaux de neurones, arbre de décision et ceux à base de connaissances, peuvent être plus adaptés pour gérer ces données complexes. Compte tenu de la variété des

algorithmes disponibles, la combinaison de différentes approches de classification a montré son utilité pour améliorer la précision de la classification.

Approche de classification	Caractéristiques	Exemple de classificateur	Résolution spatiale	Exemple de logiciel
Par pixel	- Chaque pixel est supposé pur et généralement classé dans une seule catégorie de la couverture terrestre.	- Unsupervised: (k-means, ISODATA) - Supervised (MV, Parallelepiped, k-nearest Neighbors) - ANN, SVM, AD	- Basse et moyenne	- ENVI, - ERDAS - Monteverdi - Grass
Sub pixel	- Chaque pixel est considéré comme un pixel mixte ; - Cette approche tient compte de la proportion relative des classes dans la zone représentée par le pixel.	- Fuzzy classification, - Neural networks, - SMA, fuzzy-spectral mixture analysis	Basse et moyenne	- ERDAS - ENVI
Orienté objet	- Ne traite plus le pixel de manière isolée mais des groupes de pixels (objets) dans leur contexte	Segmentation d'image et technique d'analyse d'image à base d'objets	Haute et très haute résolution	- E-cognition, - ENVI - Monteverdi
Contextuelle	- Exploite les informations spatiales entre pixels voisins	Integrated algorithm of MRFs and SVM, ICM	Très Haute résolution	- Matlab

Tableau 3 : Synthèse sur les approches de classification

10. Conclusion

La classification des images satellites a fait de grands progrès au cours des dernières décennies dans les trois domaines suivants: (1) le développement et l'utilisation d'algorithmes de classification de pointe ; (2) l'utilisation de multiples attributs ; et (3) l'intégration de données auxiliaires dans les procédures de classification.

Les caractéristiques spectrales sont les informations les plus importantes pour la classification d'images. Tant que la résolution spatiale augmente, la texture ou le contexte devient un autre attribut important à considérer.

Le succès d'une classification dépend de nombreux facteurs. La disponibilité de haute qualité des images de télédétection et des données auxiliaires, la conception d'une procédure de classification appropriée, et les compétences et l'expérience de l'analyste. Malgré, l'évènement continu de nouvelles méthodes de classification au cours des dernières années, il est difficile de choisir une méthode de classification appropriée pour un but précis. Par conséquent, le présent document a fourni un guide pour choisir la méthode de classification appropriée en donnant une brève description des différentes méthodes de classification. Cette étude a mis également en évidence l'évaluation de chacune de ces méthodes de classification.

Références

Adams, J.B., Sabol, D.E., Kapos, V., Filho, R.A., Roberts, D.A., Smith, M.O. and Gillespie, A.R., 1995, *Classification of multispectral images based on fractions of endmembers: application to land cover change in the Brazilian Amazon*. Remote Sensing of Environment, 52, pp. 137–154.

Akono, A., Tonye, E., Tcheuffa, S., 1996. *Analyse texturale des images radar par la méthode des réseaux de neurones et la logique floue*. In : Actes du 4ème colloque africain sur la recherche en informatique (CARI), Libreville, Gabon, p.695-704.

Ali, S. S., Dare, P. M., and Jones, S. D., 2009. *A Comparison of Pixel- and Object Level Data Fusion Using Lidar and High-Resolution Imagery for Enhanced Classification*. In: Innovations in Remote Sensing and Photogrammetry, Berlin, Germany: Springer-Verlag, 3–17.

Baatz, M. and Schäpe, A., 2000. *Multiresolution Segmentation – An Optimization Approach for High Quality Multi-scale Image Segmentation*. Angewandte Geographische Information sverarbeitung XII, Wichmann - Verlag, Heidelberg, pp. 12-23.

Benediktsson, J. A., Palmason, J. A., Sveinsson, J. R. and Chanussot, J. 2004. *Decision level fusion in classification of hyperspectral data from urban areas*. In: Geoscience and Remote Sensing Symposium, 2004. IGARSS '04, pages 73–76.

Benediktsson, J.A. and Kanellopoulos, I., 1999, *Classification of multisource and hyperspectral data based on decision fusion*. IEEE Transactions on Geoscience and Remote Sensing, 37, pp. 1367–1377.

Benediktsson, J. A., P. H. Swain, and O. K. Esroy. 1990. *Neural network approaches versus statistical methods in classification of multisource remote sensing data*. IEEE Transactions on Geoscience and Remote Sensing 28:540–552.

Benz, U. C., Hofmann P, Willhauck G, Lingenfelder I, and Heynen M., 2004. *Multi-resolution, object-oriented*

- fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 2004; 58:239–258.
- Bezdek, J. C. 1981. *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press.
- Blaschke, T. and Strobl, J., 2001. *What's wrong with pixels ? Some recent developments interfacing remote sensing and GIS*. GeoBIT/GIS, vol. 6, p. 12-17.
- Blaschke, T., Lang, S., Lorup, E., Strobl, J. and Zeil, P., 2000. *Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications*. In: Cremers, A., Greve, K. (Eds.), *Environmental Information for Planning, Politics and the Public*, vol. 2. Metropolis Verlag, Marburg, pp. 555_570.
- Bishop, C.M., 2006. *Pattern recognition and machine learning*, first ed, Springer.
- Breiman, L. *Bagging predictors*. Mach. Learn. 1996. International Conference on Machine Learning, Bari, Italy, 3–6 July 1996, pp. 148–156.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Frances, J. and Calpe-Maravilla, J., 2006. *Composite kernels for hyperspectral image classification*. IEEE Geoscience and Remote Sensing Letters, 3(1):93– 97, Jan. 2006.
- Camps-Valls, G., and Bruzzone, L., 2005. *Kernel-based methods for hyperspectral image classification*. IEEE Trans. Geos. and Remote Sens., 43(6):1351–1362, June 2005.
- Ceamanos, X., Waske, B., Benediktsson, J. A., Chanussot, J., Fauvel, M., and J. R. Sveinsson, 2010. *A Classifier Ensemble Based on Fusion of Support Vector Machines for Classifying Hyperspectral Data*. International Journal of Image and Data Fusion, 1(4):293–307.
- Chen, K., and Chi, H., 1998. *A method of combining multiple classifiers through soft competition on different feature sets*. Neurocomputing, 20, 227–252.
- Chen, D. and D. A. Stow, 2003. *Strategies for Integrating Information from Multiple Spatial Resolutions into Land-Use/Land-Cover Classification Routines*. Photogrammetric Engineering & Remote Sensing, 69:1279–1287.
- Chitroub, S., 2010. *Classifier Combination and Score Level Fusion: Concepts and Practical Aspects*. International Journal of Image and Data Fusion, 1(2):113–135.
- Cho, S., and Kim, J., 1995. *Combining multiple neural networks by fuzzy integral for robust classification*. IEEE Transactions on Systems, Man and Cybernetics, 25, 380–384.
- Coburn, C. A., and Roberts A. C. B.. 2004. *A multiscale texture analysis procedure for improved forest stand classification*. International Journal of Remote Sensing 25:4287–4308.
- Cracknell, A.P., 1998, *Synergy in remote sensing—what's in a pixel?* International Journal of Remote Sensing, 19, pp. 2025–2047.
- De Bruin, S., and B. G. H. Gorte. 2000. *Probabilistic image classification using geological map units applied to land-cover change detection*. International Journal of Remote Sensing 21:2389–2402.
- Dubes, R. C., and Jain, A. K., 1993. *Random field models in image analysis*. Journal of Applied Statistics, 20(5):121–154.
- Dietterich, T., 1999. *An experimental comparison of three methods for constructing ensembles of decision trees: Bagging boosting and randomization*. Machine Learning, 40, 139–158.
- Doan, H.T.X. and Foody, G.M., 2007. *Increasing soft classification accuracy through the use of an ensemble of classifiers*. International Journal of Remote Sensing. vol 28, 4609–4623.
- Ducrot, D., 2005. *Méthodes d'analyse et d'interprétation d'images de télédétection multi-sources : Extraction de caractéristiques du paysage*. Habilitation, INP-Toulouse, Toulouse, France.
- Dunn, D., Higgins, W. E., and Wakeley, J., 1994. *Texture segmentation using 2-d gabor elementary functions*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(2) :130–149.
- Durieux, L. Lagabriele, E. and Nelson A., 2008. *A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data*. ISPRS Journal of Photogrammetry and Remote Sensing 63(4), 399–408.
- Ehlers, M., Klonus, S., Astrand, P. J., and P. Rosso, 2010. *Multisensor Image Fusion for Pansharpening in Remote Sensing*. International Journal of Image and Data Fusion, 1:25–45.
- El kharki, O. Ait Belaid, M. Ait Ouahman, A. Sadgal, M. & Bijaber, N., 2002. *Application de la télédétection et des systèmes d'information géographique à l'élaboration de la carte d'occupation du sol dans la province d'Oujda au Maroc*. Journal Magrébin de Physique, Vol 2, N°1.
- El Kharki, O., Ducrot, D., & Mechbouh, J., 2012a. *Cartographie automatique à partir des images satellites: Apprentissage automatique à partir d'une base de données terrain géoréférencée et d'une segmentation d'images*. Revue GéoObservateur N°20.
- El Kharki, O., 2011. *Méthodologie d'élaboration des cartes d'occupation du sol à partir des images satellites*. Revue Géosciences N°1.
- El Kharki, O. Ducrot, D., Masse, A., and Mechbouh, J., 2012b. *Une nouvelle approche automatique d'échantillonnage pour la classification dirigée des images de télédétection multi-sources et multidates*. AARSE, 29 October-2 November 2012, El Jadida, Marocco.

- El Kharki, O., Mechbouh, J., Rouchdi, M. and El Aboudi, A., 2014b. *Evaluation des performances de l'approche pixel et orientée objet avec deux images de résolution spatiales différente : Application à la cartographie de l'arganaie (région Agadir, Maroc)*. In : Colloque GIS USERS, Meknés, Tunisia, 20-21 November 2014.
- Epstein, J., K. Payne, and E. Kramer. 2002. *Techniques for mapping suburban sprawl*. Photogrammetric Engineering & Remote Sensing 68:913.
- Evrendilek, F. and Gulbeyaz, O., 2011. *Boosted decision tree classifications of land cover over Turkey integrating MODIS, climate and topographic data*. International Journal of Remote Sensing, 32(12), 3461-3483.
- Fauvel, M., Benediktsson, J. A., Chanussot, J., and Sveinsson, J. R., 2008. *Spectral and spatial classification of hyperspectral data using svms and morphological profiles*. IEEE Transactions on Geosciences and Remote Sensing, 46(11): 3804–3814.
- Fauvel, M., Chanussot, J. and Benediktsson, J.A., 2006. *Decision fusion for the classification of urban remote sensing images*. IEEE Transactions on Geosciences and Remote Sensing 44: 2828–2838.
- Fisher, P., 1997. *The pixel: a snare and a delusion*. International Journal of Remote Sensing, 18, pp. 679–685.
- Flygare, A-M., 1997. *A comparison of contextual classification methods using Landsat TM*. International Journal of Remote Sensing, 18, pp. 3835–3842.
- Foody, G.M., 2002. *Status of land cover classification accuracy assessment*. Remote Sensing of Environment, 80, pp. 185–201.
- Foody, G.M., 2004. *Thematic map comparison: evaluating the statistical significance of differences in classification accuracy*. Photogrammetric Engineering & Remote Sensing, 70, pp. 627-633.
- Foody, G.M. and Mathur, A., 2004. *Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification*. Remote Sensing of Environment 93 (1–2), 107–117.
- Foody, G.M., Boyd, D.S. and Sanchez-Hernandez, C., 2007. *Mapping a specific class with an ensemble of classifiers*. International Journal of Remote Sensing. vol 28, 1733–1746.
- Frankot, R. T., and Chellappa R. 1987. *Lognormal random-field models and their applications to radar image synthesis*. IEEE Transactions on Geoscience and Remote Sensing 25:195–207.
- Franklin, S. E., Connery D. R., and Williams J. A.. 1994. *Classification of alpine vegetation using Landsat Thematic Mapper, SPOT HRV and DEM data*. Canadian Journal of Remote Sensing 20:49.
- Freund, Y. and Schapire, R.E., 1996. *Experiments with a New Boosting Algorithm*. In: Proceedings of the International Conference on Machine Learning, Bari, Italy, 3–6 July 1996, pp. 148–156.
- Fukunaga K., 1990. *Introduction to Statistical Pattern Recognition*. Academic Press, Boston, seconde édition.
- Gahegan, M., and G. West. 1998. *The classification of complex geographic datasets: An operational comparison of artificial neural networks and decision tree classifiers*. In: Proceedings of the 3rd International Conference on GeoComputation University of Bristol, United Kingdom, 17–19 September 1998, http://www.geocomputation.org/1998/61/gc_61.htm, Accessed 19 December 2008.
- Gallego, F.J., 2004. *Remote sensing and land cover area estimation*. International Journal of Remote Sensing, 25, pp. 3019–3047.
- Gercek, D., 2004. *Improvement of image classification with the integration of topographical data*. In Proceedings of 20th ISPRS Congress, Istanbul, Turkey, 12–23 July 2004. 35-B8, 53–58.
- Giorgos, M., Jungho, I. and Caesar, O., 2011. *Support vector machines in remote sensing: A review*. ISPRS Journal of Photogrammetry and Remote Sensing, 66, pp. 247–259.
- Goetz, S. J. et al. 2003. *IKONOS imagery for resource management: tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region*. Remote Sensing of Environment 88:195.
- Gong, P., 1994. *Integrated Analysis of Spatial Data from Multiple Sources: An Overview*. Canadian Journal of Remote Sensing, 20:349–359.
- Gong, P. and Howarth, P.J., 1992. *Frequency-based contextual classification and gray-level vector reduction for land-use identification*. Photogrammetric Engineering & Remote Sensing, 58, pp. 423–437.
- Guerschman, J. P. et al., 2003. *Land cover classification in the Argentine Pampas using multitemporal Landsat TM data*. International Journal of Remote Sensing 24:3381.
- Gurney, C. M., 1980. *Threshold selection for line detection algorithms*. IEEE Transactions on Geoscience and Remote Sensing 18:204–211.
- Gurney, C. M., 1981. *The use of contextual information to improve land cover classification of digital remotely sensed data*. International Journal of Remote Sensing 2:379–388.
- Hansen, M. C., Dubayah, R. and Defries R. S., 1996. *Classification trees: An alternative to traditional land cover classifiers*. International Journal of Remote Sensing 17:1075–1081.
- Haralick, R. M., Shanmugam, M. and Dinstein, I., 1973. *Texture feature for image classification*. IEEE Transactions on Systems, Man and Cybernetics 3:610–621.
- Hashem, S. and Schmeiser, B., 1995. *Improving model accuracy using optimal linear combinations of trained*

- neural networks. IEEE Transactions on Neural Networks, 6, 792–794.
- Haslett, J. 1985. *Maximum likelihood discriminant analysis on the plane using a Markovian model of spatial context*. Pattern Recognition 18:287–296.
- Huang, C., Davis, L. S., and Townshend, J. R. G., 2002. *An assessment of support vector machines for land cover classification*. International Journal of Remote Sensing, 23(4), 725–749.
- Huang, Z. and Lees, B.G., 2004. *Combining non-parametric models for multisource predictive forest mapping*. Photogrammetric Engineering & Remote Sensing, 70, pp. 415–425.
- Hubert-Moy, L., Cotonnec, A., Le Du, L., Chardin, A. and Perez, P., 2001. *A comparison of parametric classification procedures of remotely sensed data applied on different landscape units*. Remote Sensing of Environment, 75, pp. 174–187.
- Jackson, Q., and Landgrebe, D., 2002. *Adaptive bayesian contextual classification based on Markov random fields*. IEEE Transactions on Geoscience and Remote Sensing, 40(11):2454–2463, 2002.
- Jensen, J.R., 2004. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Third edition. Upper Saddle River, NJ.: 2004. p. 526.
- Jimenez, L. O., Morales-Morell, A., and A. Creus, 1999. *Classification of Hyperdimensional Data Based on Feature and Decision Fusion Approaches Using Projection Pursuit, Majority Voting, and Neural Networks*. IEEE Transactions on Geoscience and Remote Sensing, 37:1360–1366.
- Kandrika, S., and Roy, P. S., 2008. *Land use land cover classification of Orissa using multi-temporal IRS-P6 AWiFS data: A decision tree approach*. International Journal of Applied Earth Observation and Geoinformation, 10(2), 186–193.
- Kartikyan, B., Gopalakrishna, B., Kalubarme, M.H. and Majumder, K.L., 1994. *Contextual techniques for classification of high and low resolution remote sensing data*. International Journal of Remote Sensing, 15, pp. 1037–1051.
- Keuchel, J., Naumann, S., Heiler, M. and Siegmund, A., 2003. *Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data*. Remote Sensing of Environment, 86, pp. 530–541.
- Klonus, S. and M. Ehlers, 2007. *Image Fusion Using the Ehlers Spectral Characteristics Preservation Algorithm*. GIScience & Remote Sensing, 44(2): 93–116.
- Kulkarni, A. V., and N. K. Laveen. 1976. *An optimisation approach to hierarchical classifier design*. In Proceedings of the Third International Joint Conference on Pattern Recognition, IEEE Catalogue No. 76CH1140-3C, 459–466. Piscataway, NJ: IEEE.
- Kuo, C. C. J. and Chang, T., 1993. *Texture analysis and classification with treestructured wavelet transform*. IEEE Transactions on Image Processing, 2(4) :429–441.
- Kurzynski, M. W. 1983. *The optimal strategy of a tree classifier*. Pattern Recognition 16:81–87.
- Laberte A.S., Rango A., Havstad K.M., Paris J.F., Beck R.F., McNeely R., Gonzalez A.L., 2004. *Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico*. Remote Sensing of Environment 93:198–210.
- Lam, L., and Suen, C., 1995. *Optimal combinations of pattern classifiers*. Pattern Recognition Letters, 16, 945–954.
- Lepisto, L., Kunttuand, I., Autio, J., and Visa, A., 2003. *Classification method for colored natural textures using Gabor filtering*. In Proceedings of the 12th International Conference on Image Analysis and Processing, pages 397–401.
- Lepisto, L., Kunttuand, I., Autio, J., and Visa, A., 2003. *Rock image classification using non-homogenous textures and spectral imaging*. In WSCG short papers Proceedings.
- Licciardi, G., Pacifici, F., Tuia, D., Prasad, S., West, T., Giacco, F., Inglada, J., Christophe, E., Chanussot, J., and Gamba, P., 2009. *Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRS-S data fusion contest*. IEEE Transactions on Geoscience and Remote Sensing, 47(11):3857–3865, Nov. 2009.
- Lillesand T.M. and Kiefer R.W., 2008. *Remote Sensing and Image Interpretation*. 5th edition, John Wiley and Sons, inc. USA.
- Linde, Y., Buzo, A., and Gray, R. M., 1980. *An algorithm for vector quantizer design*. IEEE Trans. Commun., pages 84–95, 1980.
- Lu, D. et al., 2008. *A comparative study of Landsat TM and SPOT HRG images for vegetation classification in the Brazilian Amazon*. Photogrammetric Engineering & Remote Sensing 70:311.
- Lu, D. and Weng Q., 2007. *A Survey of Image Classification Methods and Techniques for Improving Classification Performance*. International Journal of Remote Sensing, 28:823–870.
- Lu, D., and Weng, Q., 2006. *Use of impervious surface in urban land use classification*. Remote Sensing of Environment 102:146.
- Lu, D., and Weng, Q., 2005. *Urban classification using full spectral information of Landsat ETM+ imagery in Marion County, Indiana*. Photogrammetric Engineering & Remote Sensing 71:1275.
- Lu, D., Moran, E. and Batistella, M., 2003. *Linear mixture model applied to Amazonian vegetation classification*. Remote Sensing of Environment, 87, pp. 456–469.
- Lucas, R. M., Cronin, N., Moghaddam, M., Lee, A., Armston, J., Bunting, P., and C. Witte, 2006. *Integration*

- of Radar and Landsat-Derived Foliage Projected Cover for Woody Growth Mapping, Queensland, Australia. *Remote Sensing of Environment*, 100:388–406.
- Lunetta, R. S., and Balogh, M. E., 1999. *Application of multi-temporal Landsat 5 TM imagery for wet land identification*. *Photogrammetric Engineering & Remote Sensing* 65:1303.
- Lunetta, R.S., Ediriwckrema, J., Iames, J., Johnson, D.M., Lyon, J.G., Mckerrow, A. and Pilant, A., 2003. *A quantitative assessment of a combined spectral and GIS rulebased land-cover classification in the Neuse river basin of North Carolina*. *Photogrammetric Engineering & Remote Sensing*, 69, pp. 299–310.
- Magnussen, S., Boudewyn, P., and Wulder, M., 2004. *Contextual classification of Landsat TM images to forest inventory cover types*. *International Journal of Remote Sensing* 25:2421–2440.
- Mallinis G, Koutsias N, Tsakiri-Strati M, Karteris M., 2008. *Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site*. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2008; 63:237–250.
- Mandelbrot, B. B. 1977. *Fractals: Form, chance and dimension*. San Francisco. CA: Freeman.
- Mandelbrot, B. B. 1982. *The fractal geometry of nature*. San Francisco, CA: Freeman.
- Mandler, E., and Schurmann, J., 1988. *Combining the classification results of independent classifiers based on the Dempster–Shafer theory of evidence*. In Gelsema, E., & Kanal, L. (Eds.), *Pattern recognition and artificial intelligence*, pp. 381–393. North-Holland.
- Maselli, F., Rodolfi, A. and Conese, C., 1996. *Fuzzy classification of spatially degraded Thematic Mapper data for the estimation of sub-pixel components*. *International Journal of Remote Sensing*, 17, pp. 537–551.
- Melgani, F., and Bruzzone, L., 2004. *Classification of hyperspectral remote sensing images with support vector machines*. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1778–1790.
- Muchoney, D., Borak, J., Chi, H., Friedl, M., Gopal, S., Hodges, J., Morrow, N., and Strahler, A. H., 2000. *Application of MODIS global supervised classification model to vegetation and land cover mapping in Central America*. *International Journal of Remote Sensing* 21:1115–1138.
- Ndi Nyoungui, A., Tony, E. and Akono, A., 1997. *Land cover classification in SAR images by means of a backpropagation neural network classifier and the fuzzy c-means clustering algorithm*. In: *Proceedings of the Geomatics in the Era of Radarsat (GER'97) Symposium*, Ottawa, Canada, 25-30 May 1997.
- Nishii, R., and S. Eguchi. 2005. *Supervised image classification by contextual AdaBoost based on posteriors in neighbourhoods*. *IEEE Transactions on Geoscience and Remote Sensing* 43:2547–2554.
- Oetter, D. R. et al., 2000. *Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data*. *Remote Sensing of Environment* 76:139.
- Otukei, J.R. and Blaschke, T., 2010. *Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms*. *International Journal of Applied Earth Observation and Geoinformation*, 12, S27-S31.
- Ouma, Y. O., Ngigi, T. G., and Tateishi R., 2006. *On the optimization and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application towards urban-tree delineation*. *International Journal of Remote Sensing* 27:73–104.
- Pal, M., and Mather, P. M., 2005. *Support vector machines for classification in remote sensing*. *International Journal of Remote Sensing* 26:1007–1011.
- Pal, M. and Mather, P.M., 2003. *An assessment of the effectiveness of decision tree methods for land cover classification*. *Remote Sensing of Environment*, 86, pp. 554–565.
- Palmason, J. A., Benediktsson, J. A., Sveinsson, J. R., and Chanussot, J., 2005. *Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis*. In *International Geoscience and Remote Sensing Symposium*, volume 1, pages 176–179.
- Pedroni L., 2003. *Improved classification of Landsat Thematic Mapper data using modified prior probabilities in large and complex landscapes*. *International Journal of Remote Sensing*.
- Pohl, C. and J. L. van Genderen, 1998. *Multisensor Image Fusion in Remote Sensing: Concepts, Methods, and Applications*. *International Journal of Remote Sensing*, 19:823–854.
- Pony, O., Descombes, X. and Zerubia, J., 2000. *Classification d'images satellitaires hyperspectrales en zone rurale et périurbaine*. Technical report, Sophia Antipolis, France, September 2000.
- Rashed, T., Weeks, J.R., Gadalla, M.S. and Hill, A.G., 2001. *Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: a case study of the Greater Cairo region, Egypt*. *Geocarto International*, 16, pp. 5–15.
- Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G. and Green, R.O., 1998, *Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models*. *Remote Sensing of Environment*, 65, pp. 267–279.
- Rogova, G., 1994. *Combining the results of several neural network classifiers*. *Neural Networks*, 7, 777–781.
- Roli, F. and Fumera, G., 2001. *Support vector machines for remote sensing image classification*. *European Remote Sensing*, pp. 160–166)
- Roussel, G., 2012. *Développement et évaluation de nouvelles méthodes de classification spatiale-spectrale*

- d'images hyperspectrales*. PhD thesis Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France.
- Rumelhart, D.E., Clelland, J.L. and Group, P.R., 1986. *Parallel distributed processing*. Exploration in the microstructure of cognition, I, II and III, A Bradford Book, Cambridge (MA), MIT Press.
- Safavian, S. R., and Landgrebe, D., 1991. *A survey of decision tree classifier methodology*. IEEE Transactions on Systems, Man and Cybernetics 21:660–674.
- Smara, Y., Ouarab, N., Laama, S., and Cherifi, D., 2003. *Techniques de fusion et de classification floue d'images satellitaires multisources pour la caractérisation et le suivi de l'extension du tissu urbain de la région d'Alger (Algérie)*. Second FIG Regional Conference Marrakech, Morocco, December 2-5.
- Shalan, M.A., Arora, M.K. and Ghosh, S.K., 2003. *An evaluation of fuzzy classifications from IRS 1C LISS III imagery: a case study*. International Journal of Remote Sensing, 24, pp. 3179–3186.
- Sharma, K.M.S. and Sarkar, A., 1998. *A modified contextual classification technique for remote sensing data*. Photogrammetric Engineering & Remote Sensing, 64, pp. 273–280.
- Sonka, M., Hlavac, V. and Boyle R., 1998. *Image processing: Analysis and Machine vision*. PWS Publishing, 770 p.
- Solaiman, B., Mouchot M. C., and Hillion, A., 1996. *Contextual dynamic neural networks learning in multispectral images classification*. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'96), Lincoln, Nebraska, May 27–31, 1996, 523–525. Piscataway, NJ: IEEE.
- Swain, P. H., and Hauska, H., 1977. *The decision tree classifier: Design and potential*. IEEE Transactions on Geoscience and Remote Sensing 15:142–147.
- Steele, B.M., 2000. *Combining multiple classifiers: an application using spatial and remotely sensed information for land cover type mapping*. Remote Sensing of Environment, 74, pp. 545–556.
- Stow D., Lopez A., Lippitt C., Hinton S., Weeks J., 2007. *Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data*. International Journal of Remote Sensing. 2007; 28:5167–5173.
- Stuckens, J., Coppin, P.R. and Bauer, M.E., 2000, *Integrating contextual information with per-pixel classification for improved land cover classification*. Remote Sensing of Environment, 71, pp. 282–296.
- Tarabalka, Y., Benediktsson, J. A., and Chanussot, J., 2009. *Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques*. IEEE Transactions on Geoscience and Remote Sensing, 47(8) :2973–2987.
- Tarabalka, Y., Chanussot, J., and Benediktsson, J. A., 2010a. *Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers*. IEEE Transactions on Systems, Man, and Cybernetics, 40(5) :1267–1279.
- Theodoridis, S., Pikrakis, A., Koutroumbas, K., Cavouras, D., 2010. *Introduction to Pattern Recognition: A Matlab Approach*. Edition Elsevier.
- Thomas N., Hendrix C., Congalton R.G., 2003. *A comparison of urban mapping methods using high-resolution digital imagery*. Photogrammetric Engineering & Remote Sensing. 2003; 69:963–972.
- Tonye, E. and Akono, A., 2000. *Le traitement des images de télédétection par l'exemple*. Gordon and Breach Science Publishers.
- Tso, B. and Mather, P., 2009. *Classification Methods for Remotely Sensed Data*. Second edition, CRC Press Inc, Taylor & Francis Group, USA.
- Tso, B. and Olsen, R. C., 2005. *Combining spectral and spatial information into hidden Markov models for unsupervised image classification*. International Journal of Remote Sensing 26:2113–2133.
- Vapnik, V., 1998. *Statistical learning theory*. Wiley, New York.
- Wang, F., 1990. *Fuzzy Supervised Classification of Remote Sensing Images*. IEEE Transactions on Geoscience and Remote Sensing 28(2).
- Waske, B. and Benediktsson, J.A., 2007. *Fusion of support vector machines for classification of multisensory data*. IEEE Transactions on Geoscience and Remote Sensing, 45, 3858–3866.
- Wen, X., Hu, G. and Yang, X., 2008. *CBERS-02 remote sensing data mining using decision tree algorithm*. In International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia and Workshop, Article No. 58.
- Wolter, P. T. et al., 1995. *Improved forest classification in the northern lake states using multi-temporal Landsat imagery*. Photogrammetric Engineering & Remote Sensing, 61:1129.
- Wong, T. H., Mansor, S. B., Mispan, M. R., Ahmad, N., Sulaiman, W. N. A., 2003. *Feature extraction based on object oriented analysis*. In: Proceedings of ATC Conference, 20-21 May 2003, Kuala Lumpur, Malaysia.
- Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., Schirokauer, D., 2006. *Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery*. Photogrammetric Engineering & Remote Sensing, 72:799–811.
- Zadeh, L. A. 1965. *Fuzzy sets*. Information Control 8:338–353.
- Zhang, J., 2010. *Multisource Remote Sensing Data Fusion: Status and Trends*. International Journal of Image and Data Fusion, 1:5–24.
- Zhou, Q. and Robson, M., 2001. *Contextual information is ultimately necessary if one is to obtain accurate*

image classifications. International Journal of Remote Sensing, 22, pp. 3457-3470.

Zhang, J. and Foody, G.M., 2001. *Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical neural network approaches*. International Journal of Remote Sensing, 22, pp. 615–628.

Zhang, J. and Kirby, R.P., 1999. *Alternative criteria for*

defining fuzzy boundaries based on fuzzy classification of aerial photographs and satellite images. Photogrammetric Engineering & Remote Sensing, 65, pp. 1379–1387.

Zhou, W., Troy, A., and Grove J. M., 2008. *Object-based land cover classification and change analysis in the Baltimore metropolitan area using multi-temporal high resolution remote sensing data*. Sensors 8:1613.