



Laboratoire des sciences de l'image,
de l'informatique et de la télédétection
CNRS/ULP UMR 7005



Apprentissage et Fouille de Données



Sélection automatique d'attributs et classification d'objets complexes

Alexandre Blansché
Pierre Gançarski

En collaboration avec
le Laboratoire Image et Ville
CNRS/ULP UMR 7011

Plan

- Contexte
- Notre méthode
- Application
- Conclusion

Contexte

Classification non supervisée d'objets complexes, c'est-à-dire des objets représentés par des attributs nombreux, de types variés (valeur numérique ou symbolique, histogramme, etc.)

Problèmes : attributs corrélés, bruités, non pertinents

Solution : pondération des attributs

Problèmes : l'importance des attributs dépend de la classe à mettre en évidence

Solution : rechercher une pondération par classe à mettre en évidence

Contexte

Exemple

Classification de régions obtenues par la segmentation d'une image de télédétection

Régions décrites par de nombreux attributs, de types hétérogènes :

- histogramme des classes radiométriques
- pourcentage de pixels de la frontière
- rapport entre la surface de la région et le poly. englobant
- rapport entre la surface de la région et le cercle englobant
- histogramme des classes des voisins
- moyennes radiométriques
- taille de la région
- nombre moyen de pixels voisins mis dans la même région

Pour la classe d'eau, la radiométrie est primordiale, pour le bati, la distinction se fait principalement sur la texture, pour les bâtiments (à une résolution plus élevée), c'est la forme qui est importante

Notre méthode

- Utilisation d'un classifieur par classe d'objets cherchée
- Chaque classifieur utilise une pondération différente lui permettant de mettre au mieux en évidence sa classe

 Problème : comment choisir cet ensemble de pondérations, sans l'aide d'un expert, pour que l'ensemble des classes forme une bonne classification

L'apprentissage va consister à trouver de façon non supervisée cet ensemble

Notre méthode

Apprentissage

- Impossible de faire évoluer les classifieurs de façon indépendante
⇒ nos classifieurs vont évoluer les uns en fonction des autres
- Initialisation aléatoire des pondérations
- Il faut définir pour l'apprentissage
 - Comment choisir une classe dans un classifieur
 - Quel critère pour définir la qualité d'une classification
 - Comment faire évoluer les classifieurs

Notre méthode

Choix d'une classe dans un classifieur

La classe associée à un classifieur est celle qui maximise un critère de qualité (par exemple la compacité)

- Permet de faire la correspondance entre une pondération et une classe
- Assure la qualité des classes dans le résultat final

Notre méthode

Qualité d'une classification

Les classes obtenues doivent former un partitionnement de l'ensemble des données :

Des objets non classifiés ou des objets classifiés plusieurs fois doivent faire baisser la qualité

Risque d'avoir une classe qui englobe la quasi-totalité de l'image et que la classification forme ainsi une partition de l'espace des données

Critère sur la taille des classes :

les classes doivent être de tailles comparables

Notre méthode

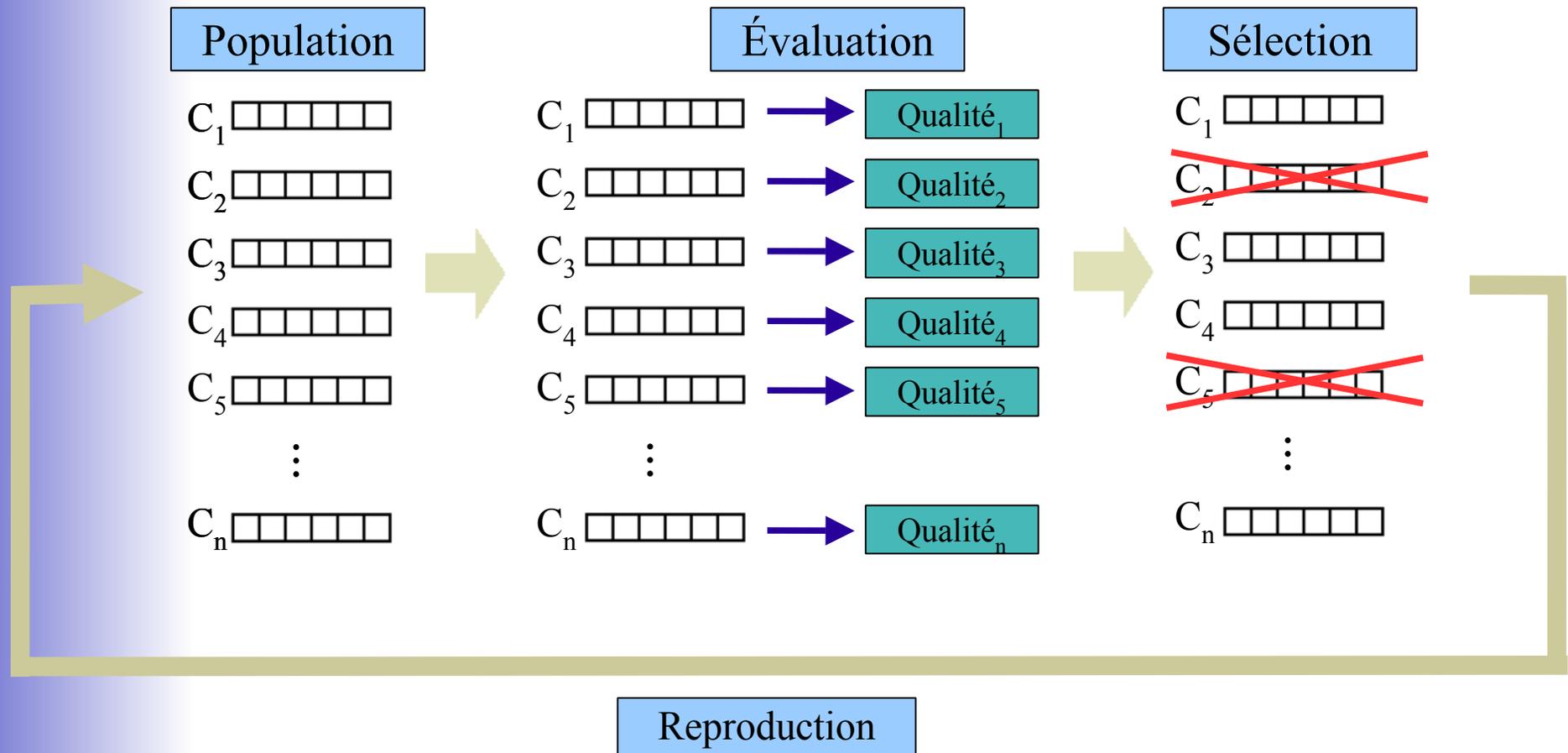
Évolution

Apprentissage par algorithme génétique

- ↳ La pondération de chaque caractéristique pour un classifieur est un gène
Un chromosome est l'ensemble des pondérations d'un ensemble de classifieur

Notre méthode

Algorithme génétiques : principes généraux



Notre méthode

Évolution

S'il y a m classes cherchées et n caractéristiques, il y aura $m \times n$ gènes, or les méthodes classiques sont peu efficaces lorsque l'espace de recherche est grand

↳ Approche par coévolution : chaque individu est un classifieur et va évoluer en fonction des autres

Notre méthode

Coévolution

- Génération d'un ensemble de classifieurs $\{C_1, \dots, C_N\}$, chaque classifieur est représenté par un vecteur des pondérations des caractéristiques
- Pour chaque classifieur, sélection de sa classe associée, en fonction d'un critère de compacité
 - ↳ Obtention des classes $\{c_1, \dots, c_N\}$
- Recherche des ensembles $\{c_{i1}, \dots, c_{im}\} \subset \{c_1, \dots, c_N\}$ qui forment les meilleures classifications de l'image, m étant le nombre de classes cherchées
- Évolution génétique pour améliorer le résultat

Notre méthode

Coévolution

- Création de m populations P_1, \dots, P_m de classifieurs (une population par classe cherchée), chaque population contenant p classifieurs et donc p classes associées
- Création de toutes les répartitions formée chacune en prenant une classe associée de chaque population
- Calcul de la qualité des répartitions, afin de trouver la meilleure
- Fonction de qualité d'un classifieur définie par la meilleure répartition à laquelle il a participé

Notre méthode

Problèmes

m le nombre de populations

p le nombre d'individus dans une population

↳ p^m répartitions possibles

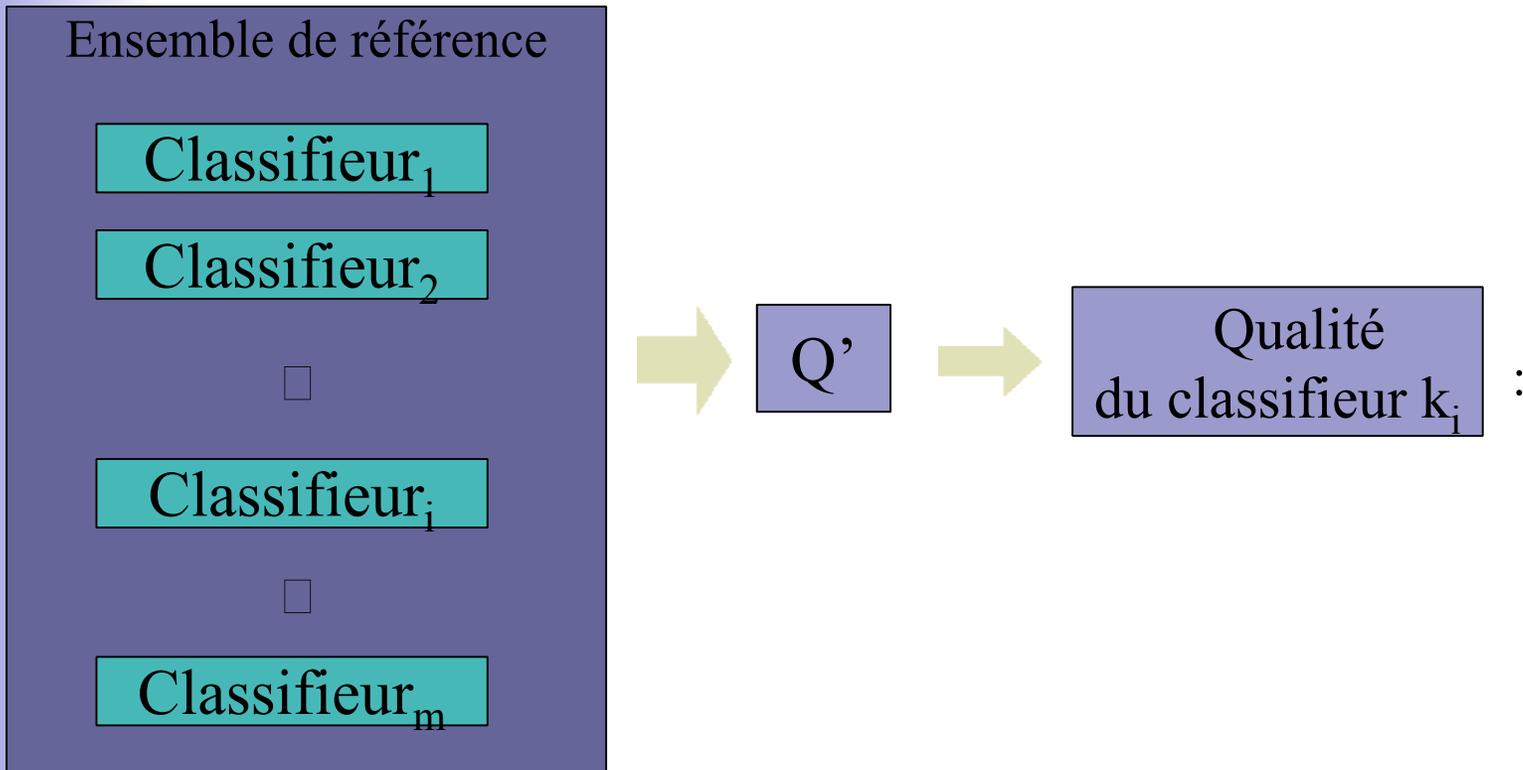
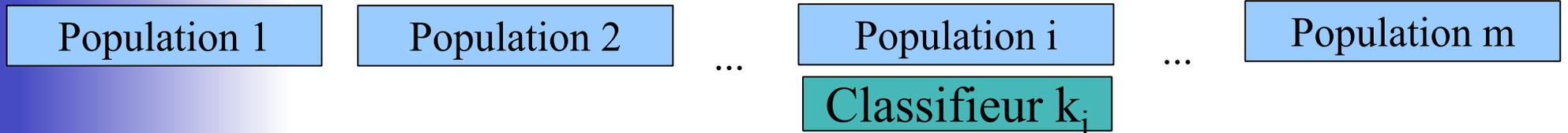
Notre méthode

Coévolution

- Fonction de qualité d'un classifieur définie en utilisant la meilleure répartition retenue dans les générations précédentes
- Évolution de la meilleure répartition en fonction des meilleurs classifieurs de la génération actuelle

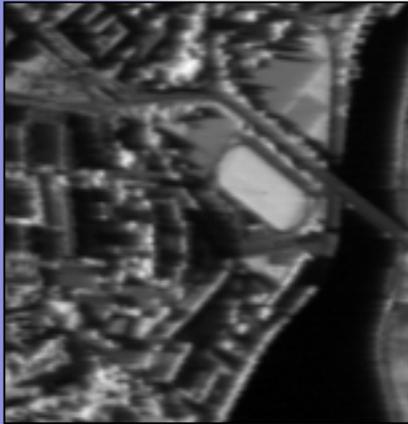
Notre méthode

Calcul de la qualité d'un classifieur



Application

Images hyperspectrales



Bande 18



Bande 24



Bande 26



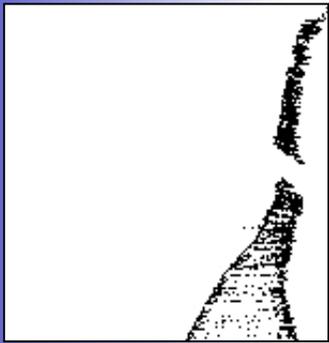
Bande 29



Bande 31

Application

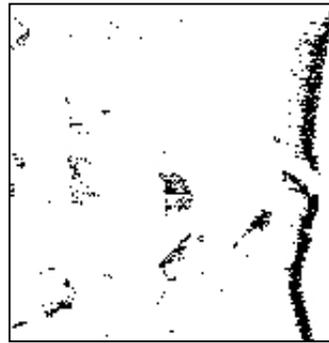
Images hyperspectrales



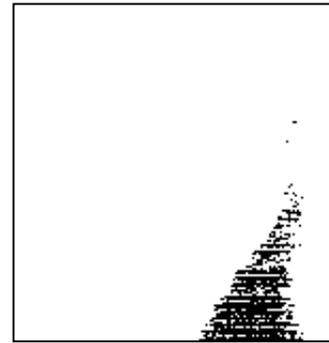
Classifieur 1



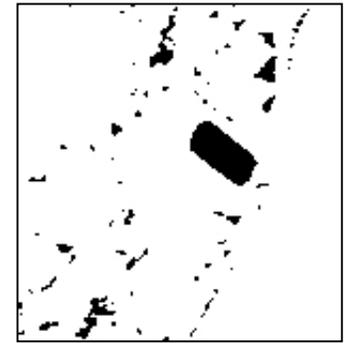
Classifieur 2



Classifieur 3



Classifieur 4

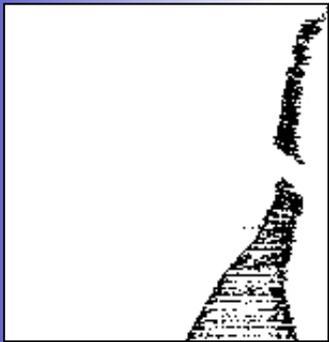


Classifieur 5

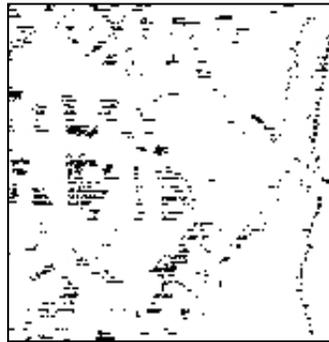
Génération 12

Application

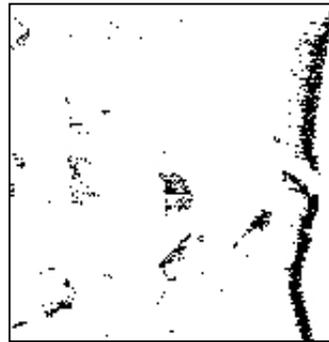
Images hyperspectrales



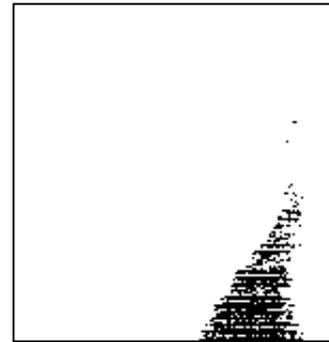
Classifieur 1



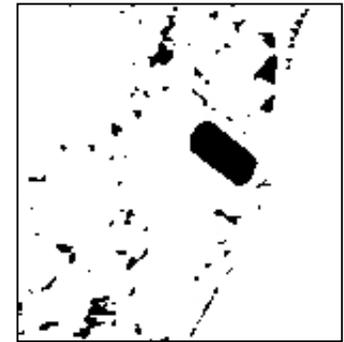
Classifieur 2



Classifieur 3



Classifieur 4



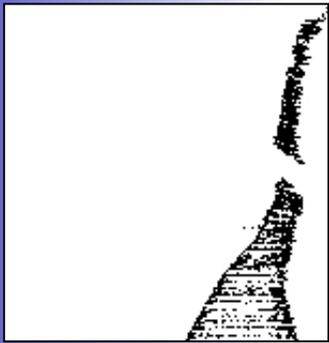
Classifieur 5

Génération 22

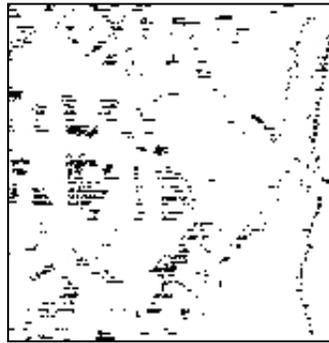
Peu d'évolution jusqu'à 28^e génération

Application

Images hyperspectrales



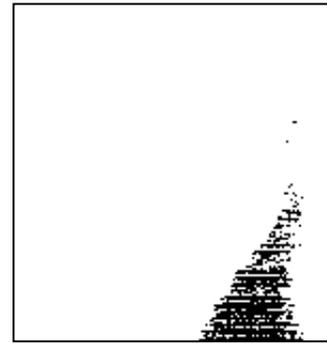
Classifieur 1



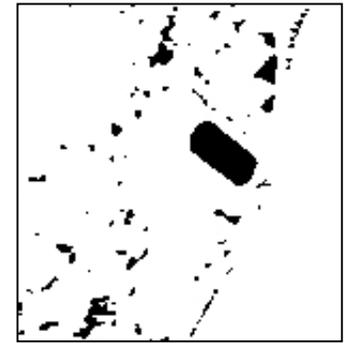
Classifieur 2



Classifieur 3



Classifieur 4



Classifieur 5

Génération 29

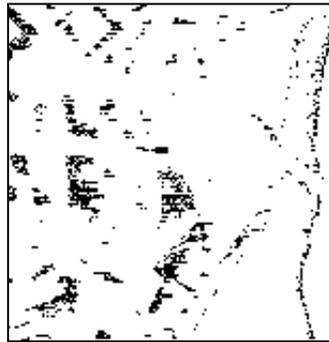
Le classifieur 3 découvre une nouvelle classe

Application

Images hyperspectrales



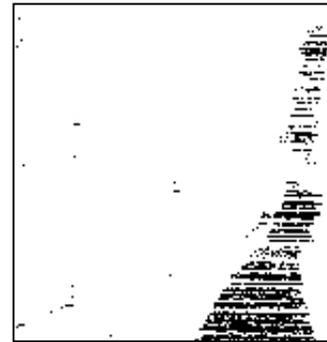
Classifieur 1



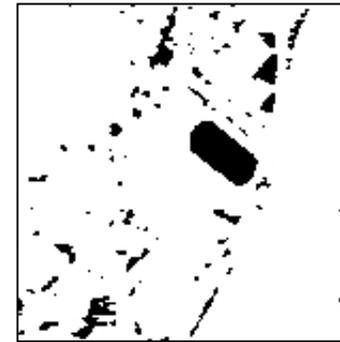
Classifieur 2



Classifieur 3



Classifieur 4



Classifieur 5

Génération 31

Le classifieur 1 découvre une nouvelle classe
Permet au classifieur 4 de changer de classe

Application

Images hyperspectrales



Classifieur 1



Classifieur 2



Classifieur 3



Classifieur 4



Classifieur 5

Génération 46

Evolution des classes

Très bonne détection de l'eau (classifieur 4), des routes (Classifieur 1) et de la verdure (classifieur 5)

Conclusion

- Nouvelle méthode de classification utilisant des classifieurs spécialisés
- Définition d'un nouveau critère de qualité d'une classification
- Apprentissage par coévolution coopérative
- Risque de convergence vers une partition correcte, mais sémantiquement incohérente \Rightarrow difficulté à trouver une fonction de qualité qui modélise au mieux le problème

Conclusion

Perspectives

- Utilisation d'une approche floue pour définir la qualité d'une classification
- Application à d'autres problèmes nécessitant la classification d'attributs hétérogènes

