

# Projet IML

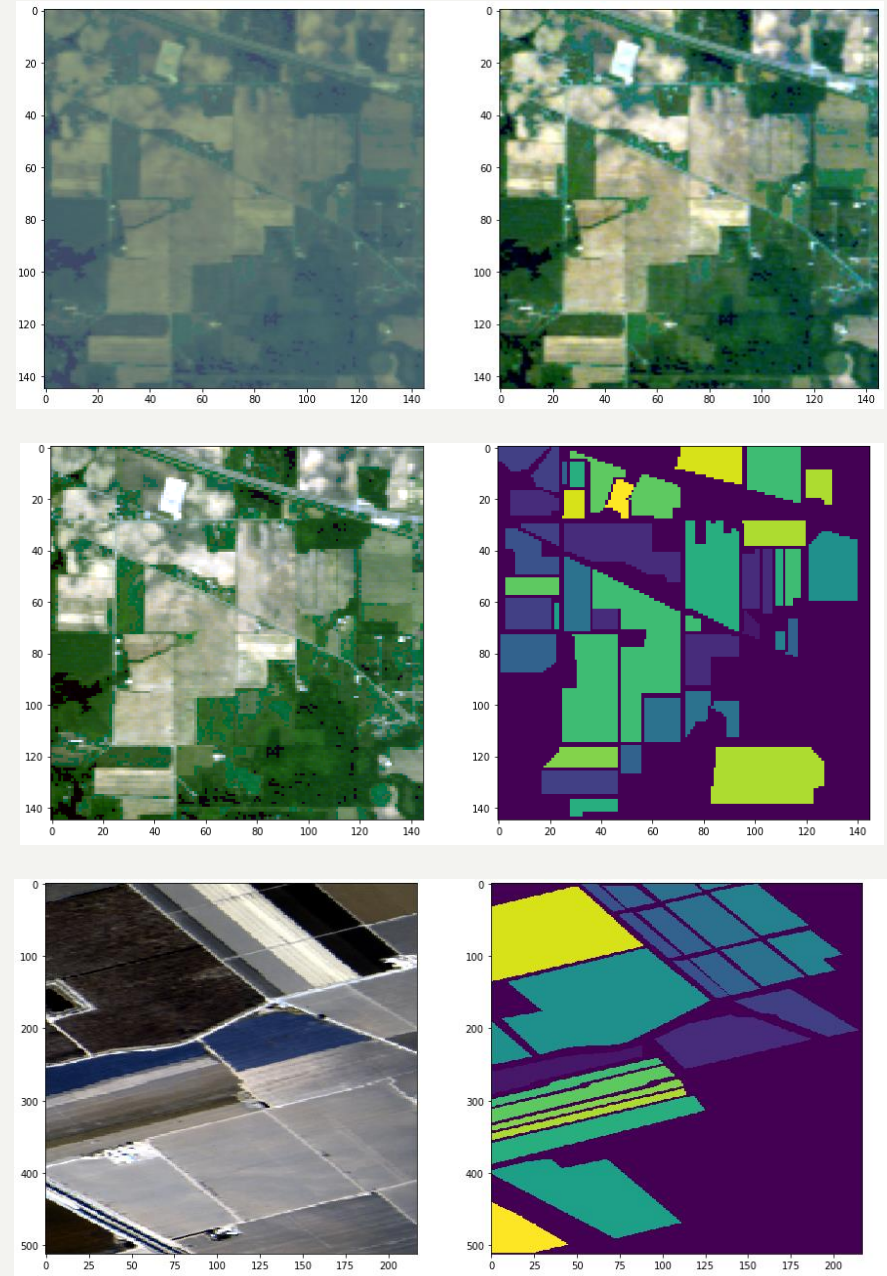
## Présentation d'une pipeline de classification d'image hyperspectrales de champs

+

Dominique Michel  
Alexandre Bourquelot  
Alexandre Pedrosa

# Le problème

- Classification des types de champs en fonction de leur spectre
- Les photos sont prises de différents angles et sur différentes bandes spectrales
- Données non-équilibrés



# Les solutions envisagées

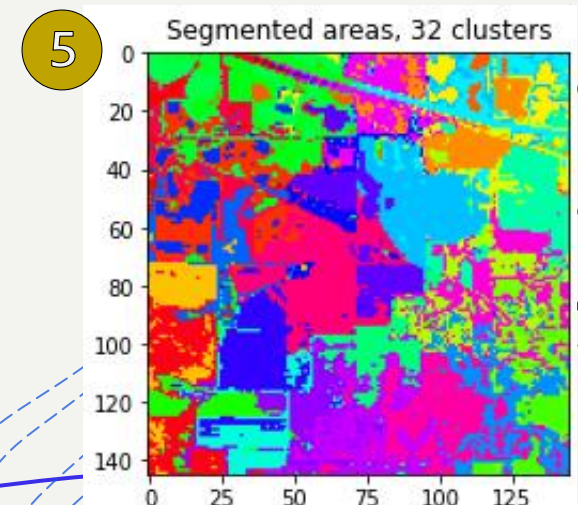
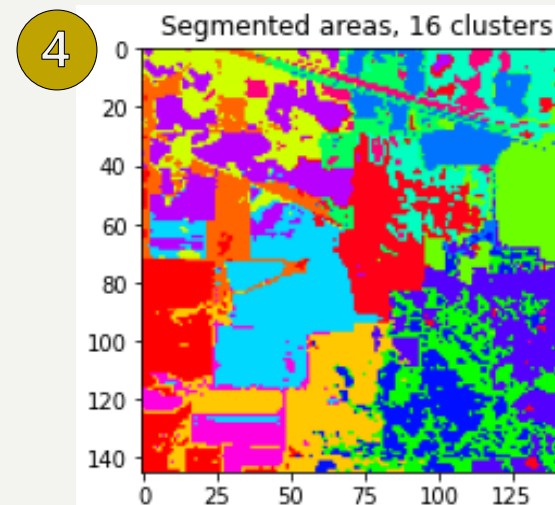
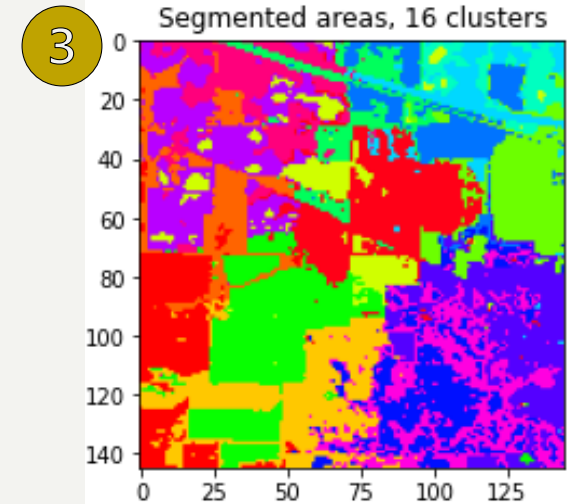
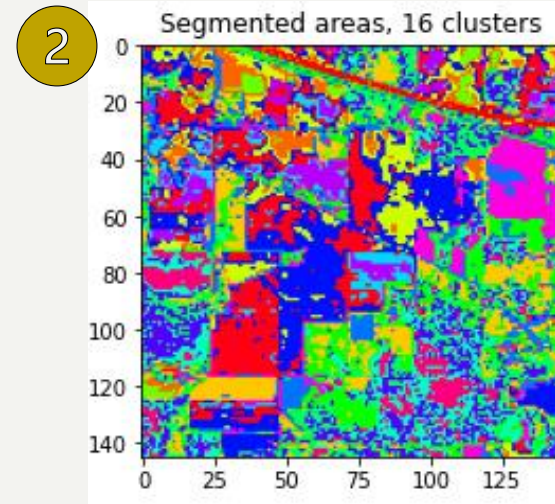
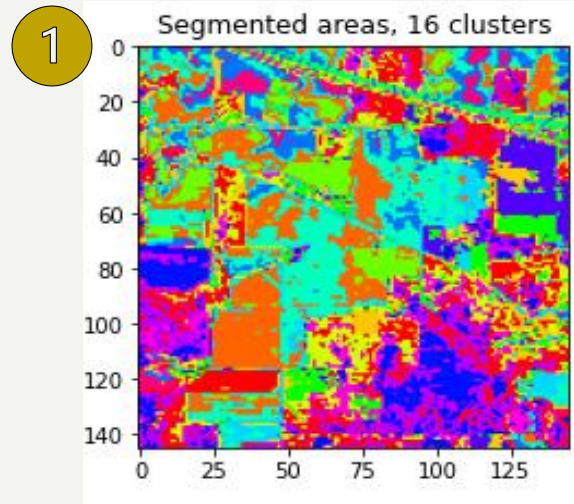
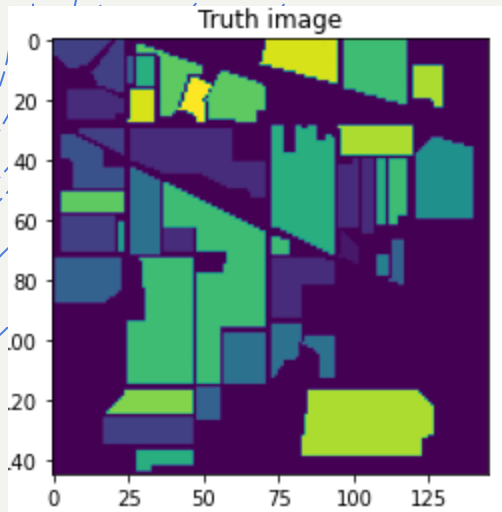
- + Réduction de dimension (200+ caractéristique)
  - + Analyse par composantes principales
- + Clustering et classification
  - + Classification des zones segmentés
  - + Classification des sous-zones par classification de la majorité des pixels
  - + Watershed par gradient sur bandes spectrales
- + Classification de chaque pixel et post-traitement
  - + Classification basé voisins, vecteurs, densité, arbre de décision
  - + Post-traitement utilisant résultat clustering ou morphologie

# Classification: Les classes testés

- +DecisionTreeClassifier
- +RandomForestClassifier
- +ExtraTreesClassifier
  
- +SVC (rbf)
- +LinearSVC
- +KNeighborsClassifier



# Clustering: Minibatch KMeans



## **Différentes informations:**

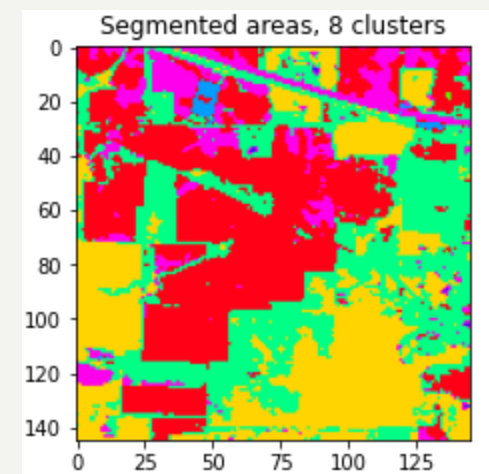
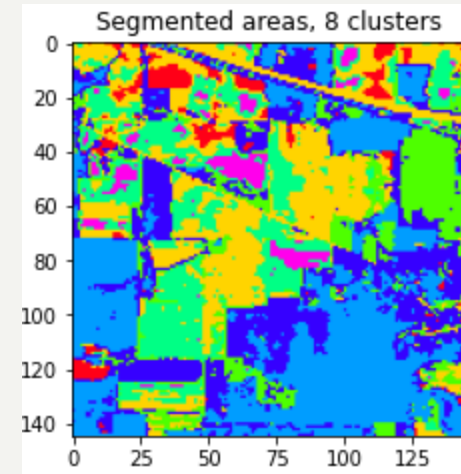
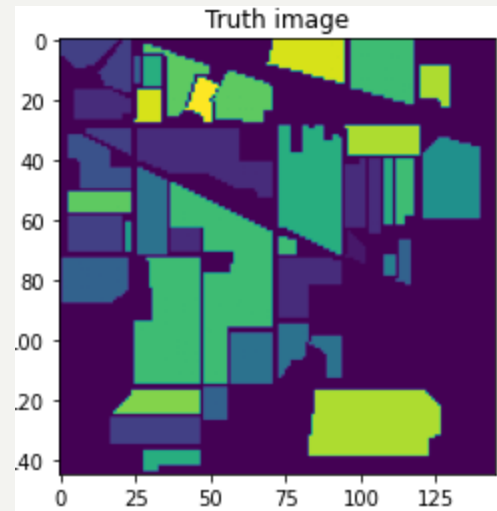
1. Sans coordonnées, sans PCA
2. Sans coordonnées
3. Sans PCA
4. Avec coordonnées, avec PCA
5. Avec coordonnées, avec PCA et 32 clusters

# Clustering: other methods

## Birch (Clustering Feature Tree)

threshold = 0.2

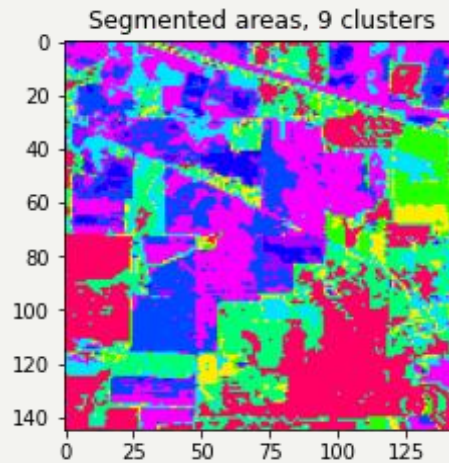
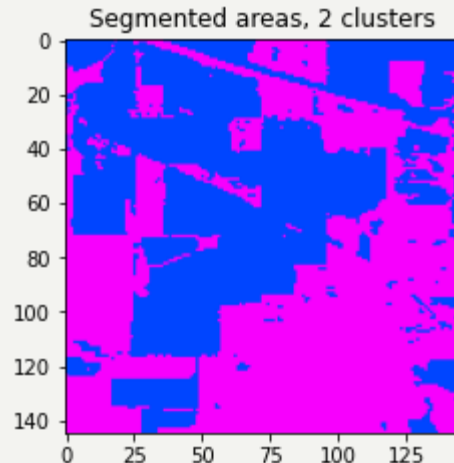
threshold = 0.3



## MeanShift (quantile)

quantile = 0.2

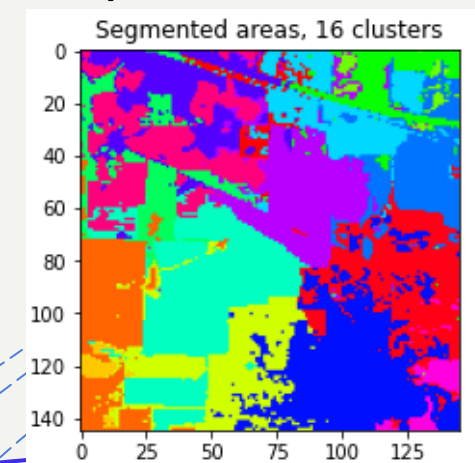
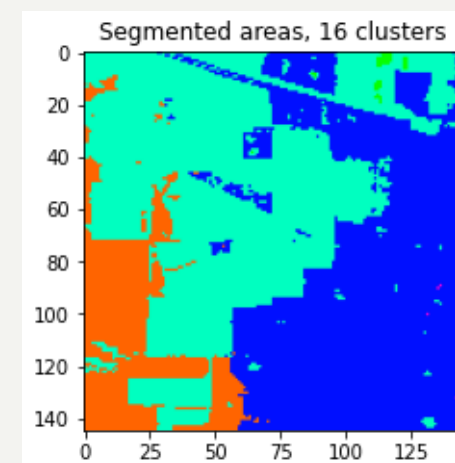
quantile = 0.01



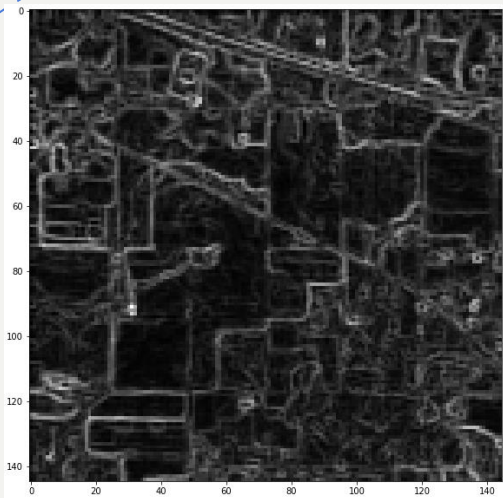
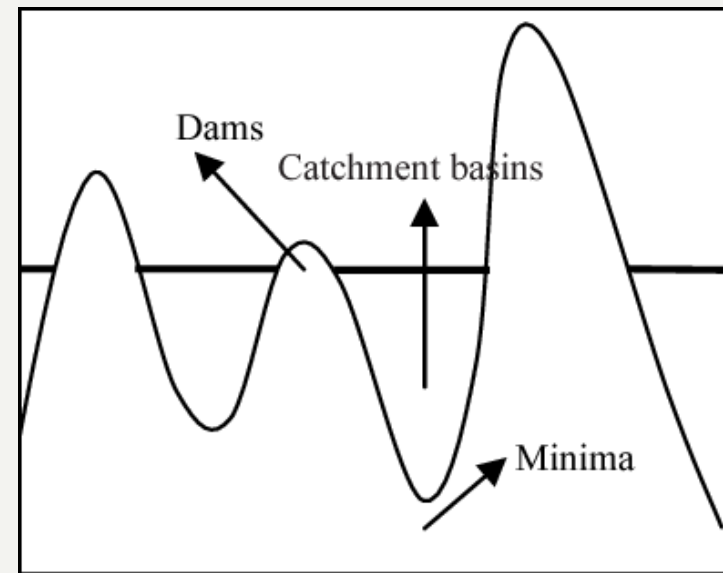
## GaussianMixture (covariance\_type)

'full' = own covariance matrix

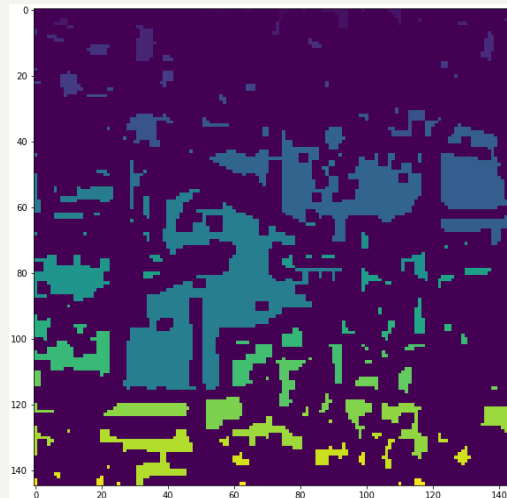
'spherical' = own variance



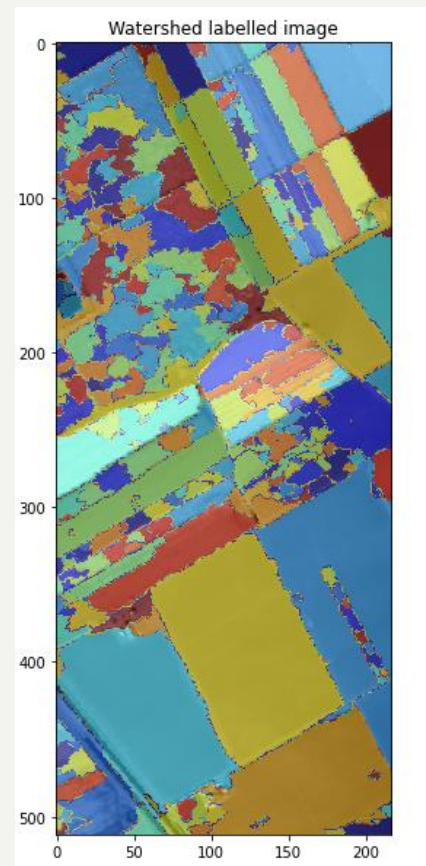
# Clustering: Watershed



Les gradients



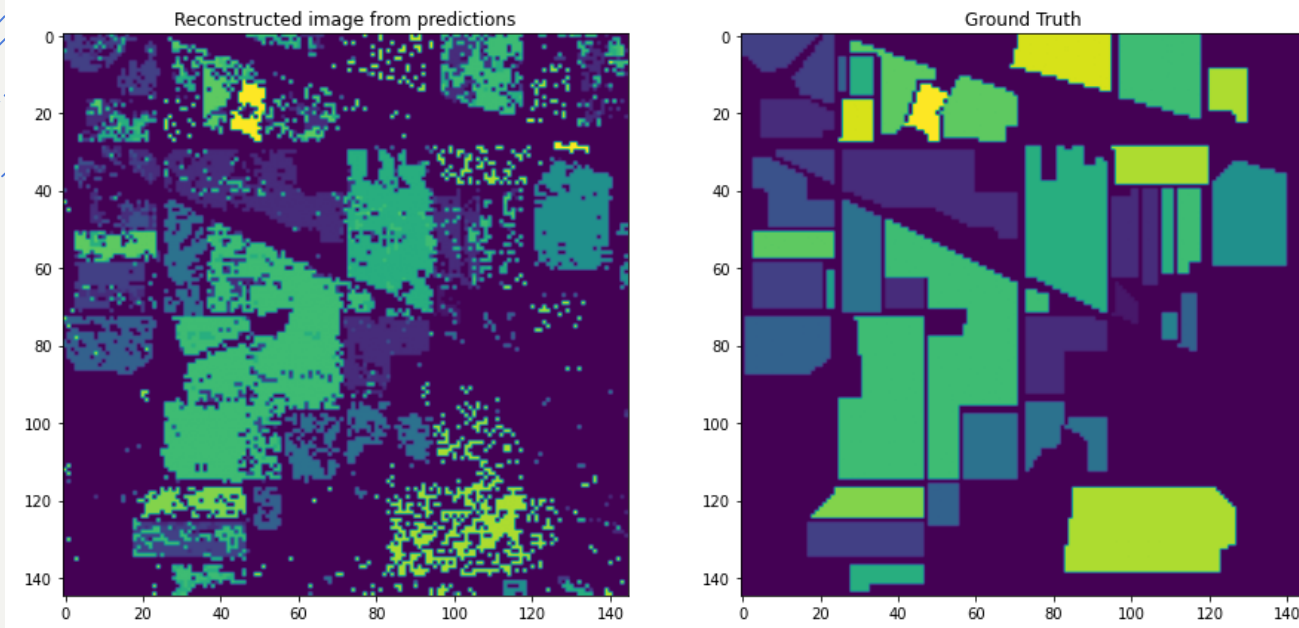
Les marqueurs





# Protocole expérimental: Comment évaluer et mesurer l'efficacité de nos solutions

Image comparison



Comparer les outputs et la vérité terrain "à l'œil nu"

	unknown	alfalfa	corn-notill	corn-mintill	corn	grass-pasture	grass-trees	grass-pasture-mowed	hay-windrowed	oats	soybean-notill	soybean-mintill	soybean-clean	wheat	woods	buildings-grass-trees-drives	stone-steel-towers
unknown	0.9	0.00011	0.019	0.0059	0.00098	0.0022	0.007	0	0.0064	0	0.011	0.019	0.0032	0.0011	0.026	0	0.0033
alfalfa	0.44	0.41	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0
corn-notill	0.22	0	0.59	0.027	0.005	0	0	0	0	0.032	0.11	0.017	0	0	0	0	0
corn-mintill	0.23	0	0.12	0.48	0.0045	0	0	0	0	0.012	0.14	0.0089	0	0	0	0	0
corn	0.59	0	0.25	0.015	0.14	0	0	0	0	0	0.01	0	0	0	0	0	0
grass-pasture	0.45	0	0	0	0	0.54	0	0	0	0	0	0	0	0	0.0098	0	0
grass-trees	0.42	0	0	0	0	0	0.58	0	0	0	0	0	0	0	0	0	0
grass-pasture-mowed	0.91	0	0	0	0	0	0.087	0	0	0	0	0	0	0	0	0	0
hay-windrowed	0.13	0	0	0	0	0	0	0.87	0	0	0	0	0	0	0	0	0
oats	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
soybean-notill	0.23	0	0.022	0.0085	0	0	0	0	0	0.61	0.11	0.018	0	0	0	0	0
soybean-mintill	0.16	0	0.064	0.025	0	0	0	0	0	0.037	0.71	0.0075	0	0	0	0	0
soybean-clean	0.42	0	0.096	0.018	0.0039	0	0	0	0	0.055	0.035	0.37	0	0	0	0	0
wheat	0.38	0	0	0	0	0	0	0	0	0	0	0	0.62	0	0	0	0
woods	0.68	0	0	0	0	0.00093	0	0	0	0	0	0	0	0.32	0	0	0
buildings-grass-trees-drives	0.96	0	0	0	0	0.0031	0.0062	0	0	0	0	0	0	0.0092	0.012	0.0062	0
stone-steel-towers	0.26	0	0.039	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7

Matrice de confusion



# Protocole expérimental: Comment évaluer et mesurer l'efficacité de nos solutions

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

	DecisionTreeClassifier	RandomForestClassifier	ExtraTreesClassifier	SVC(rbf)	LinearSVC	KNeighborsClassifier
precision	0.61	0.78	0.77	0.62	0.75	0.71
recall	0.61	0.72	0.72	0.40	0.50	0.68
f1-score	0.61	0.74	0.74	0.40	0.56	0.69
fit_duration	0.24	1.41	0.55	0.57	35.42	0.11
predict_duration	0.03	0.34	0.39	2.62	0.03	1.30

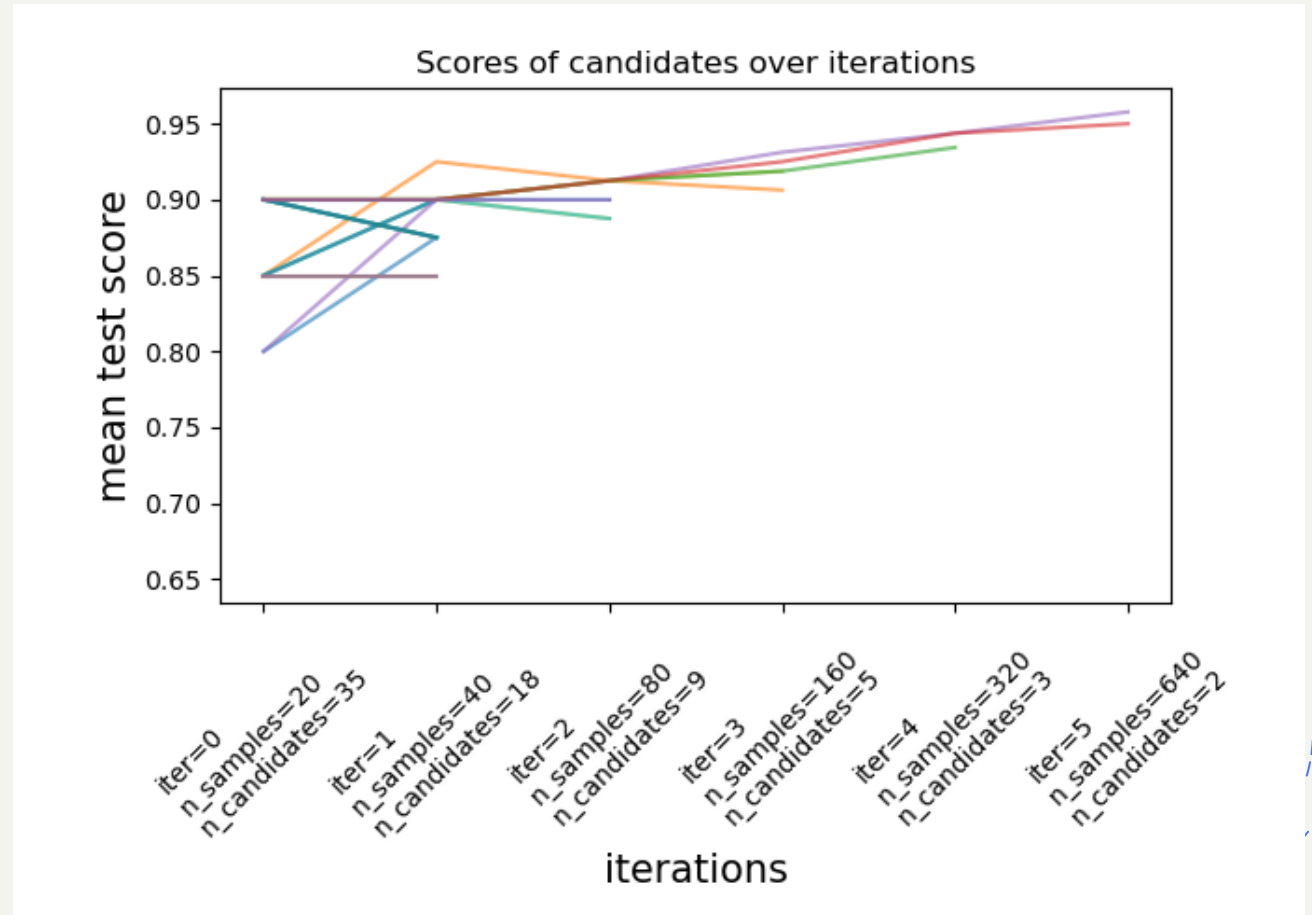
Interpréter la Précision, le Recall et le Score F1

	Before	After
precision	0.78	0.83
recall	0.74	0.78
f1-score	0.75	0.79

Observer l'effet de méthodes de pré/post processing

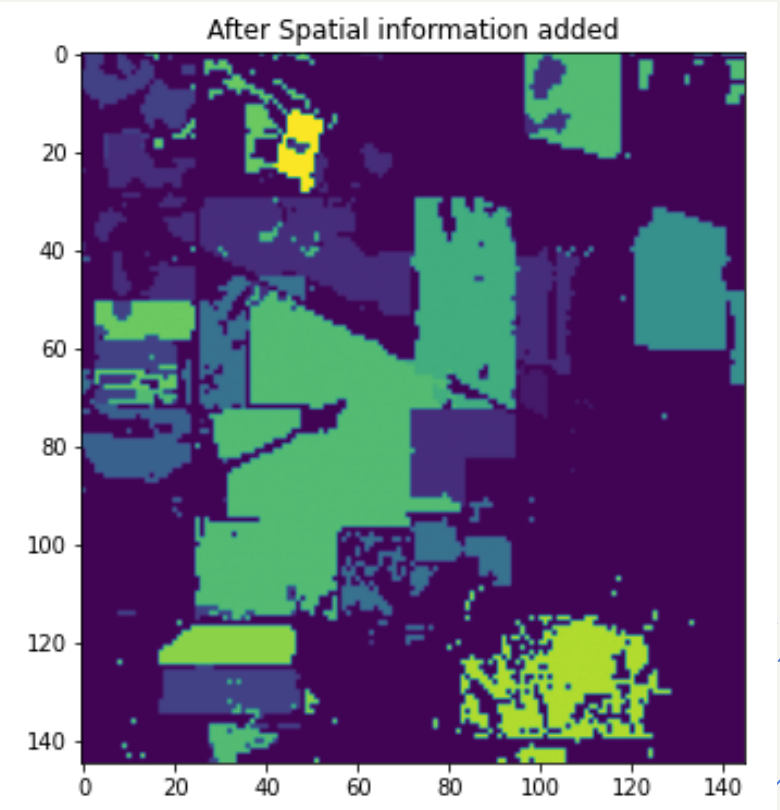
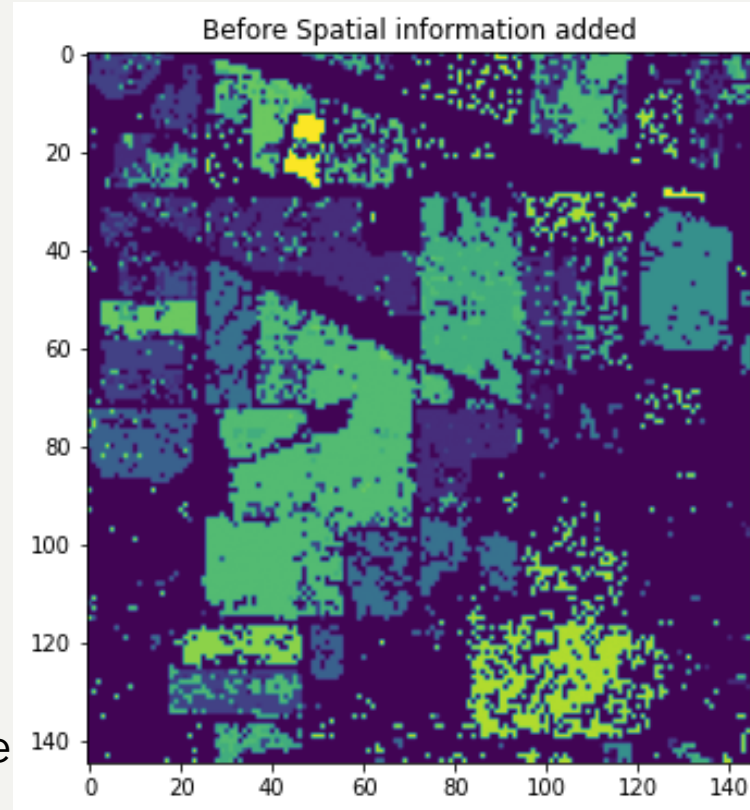
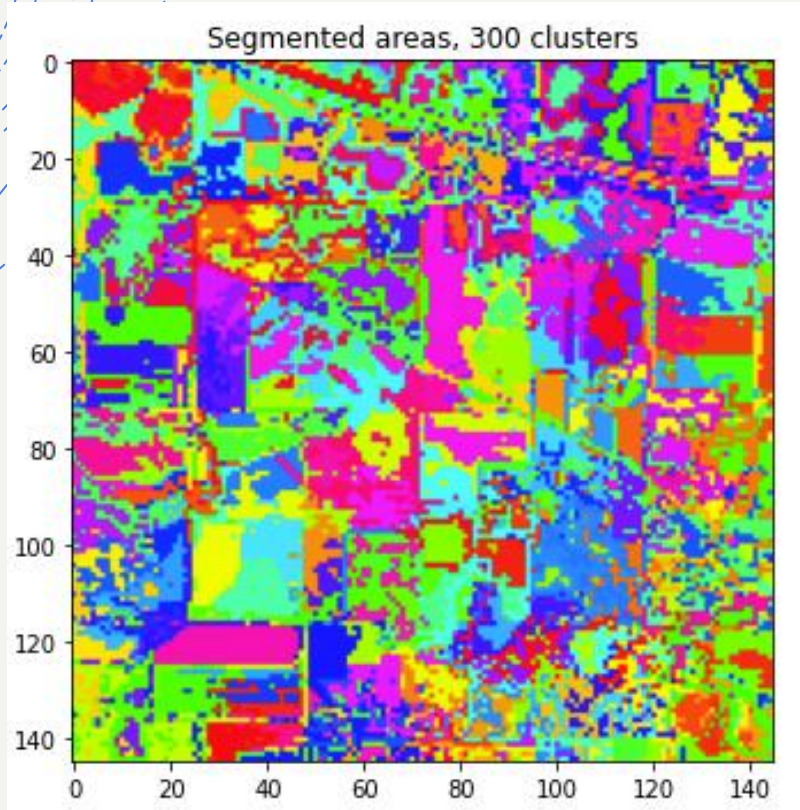
# Protocole Experimental: Grid Search

- Intégré à scikit-learn
- Automatise la recherche des paramètres optimaux



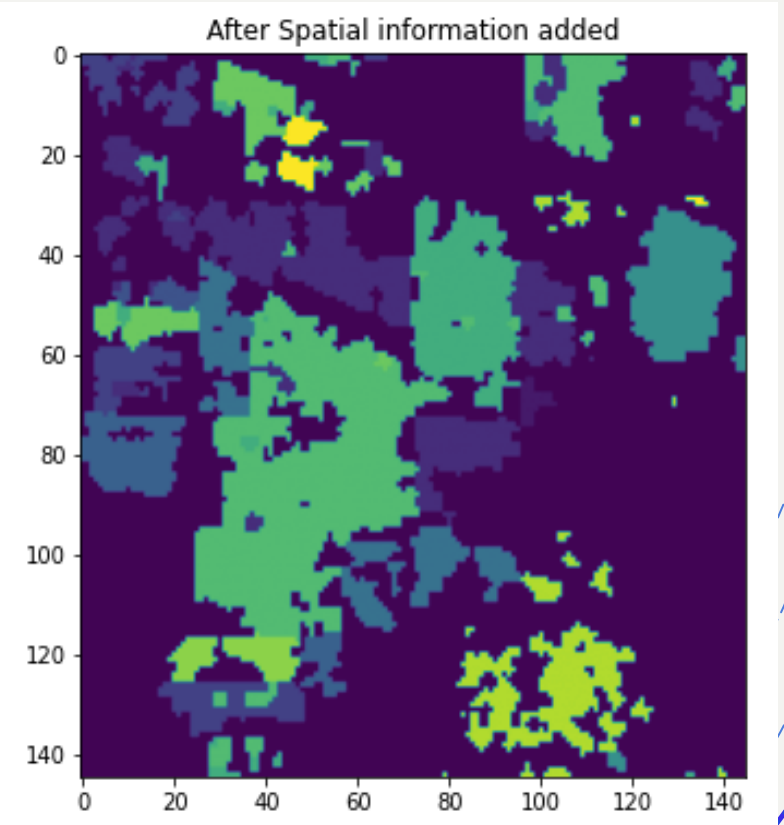
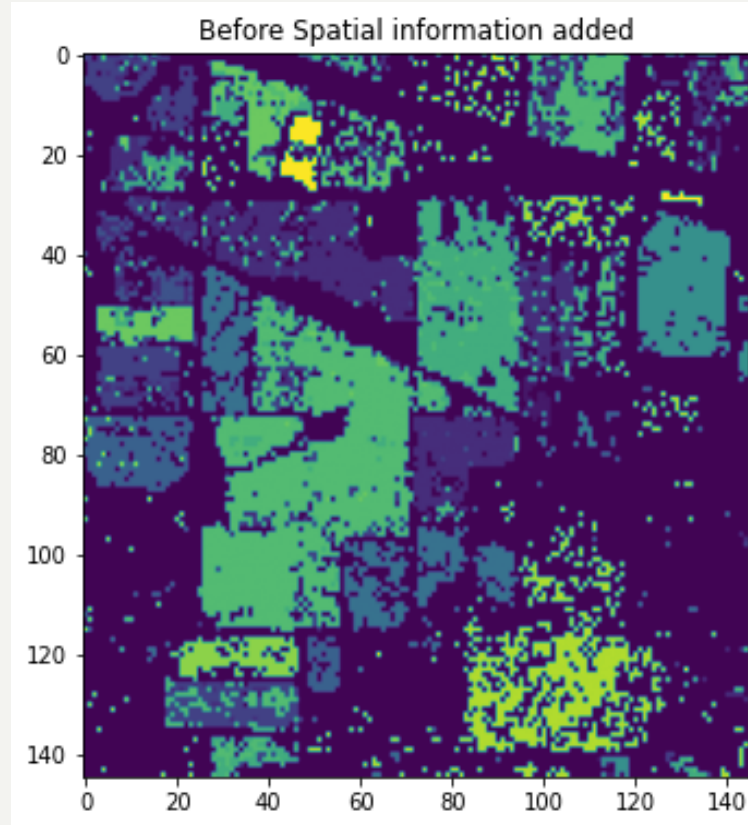
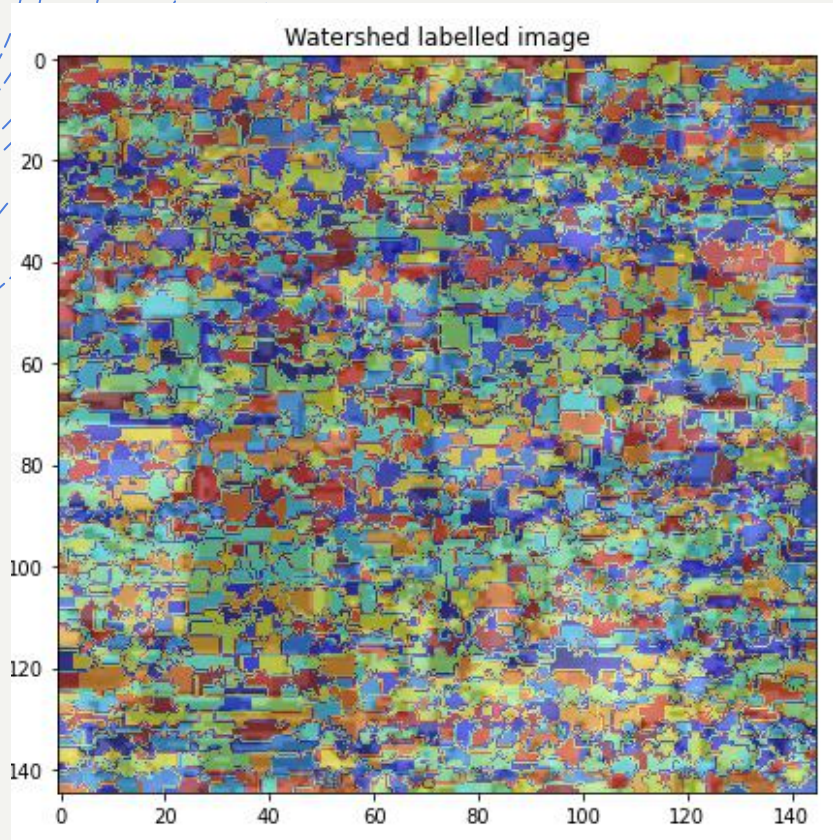
Source: scikit-learn.org

# Information Spatiale



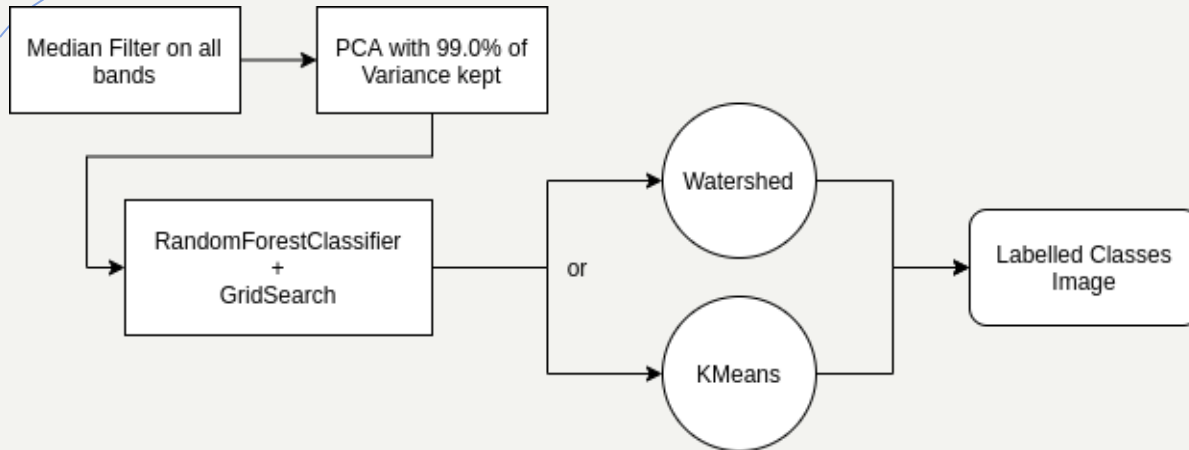


# Information Spatiale





# Résultats et Conclusion



Solution retenue:

- Pré-processing:
  - on applique un filtre médian sur les bandes
  - on applique une APC en conservant une grande partie de la variance
- Random Forest et Grid Search
- Post-Processing:
  - on ajoute de l'information spatiale