

# Apprentissage pour l'aide au diagnostic en imagerie multi-modalités du cancer

Carole Lartizien

[carole.lartizien@creatis.insa-lyon.fr](mailto:carole.lartizien@creatis.insa-lyon.fr)

CREATIS, Lyon, France



- **Place de l'imagerie médicale dans le diagnostic du cancer**
  - de plus en plus de données...
  
- **Imagerie médicale et apprentissage pour l'aide au diagnostic**
  - Principe des systèmes d'aide au diagnostic
  - Application en imagerie IRM du cancer de la prostate
  
- **Quelques enjeux de l'apprentissage pour l'aide au diagnostic**

- **Place de l'imagerie médicale dans le diagnostic du cancer**
  - de plus en plus de données...
- Imagerie médicale et apprentissage pour l'aide au diagnostic
  - Principe des systèmes d'aide au diagnostic
  - Application en imagerie IRM du cancer de la prostate
- Quelques enjeux de l'apprentissage pour l'aide au diagnostic

## Détection



IRM



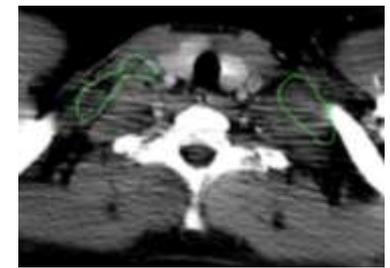
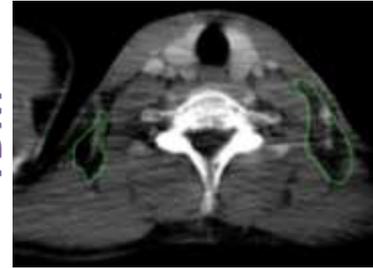
TEP FDG



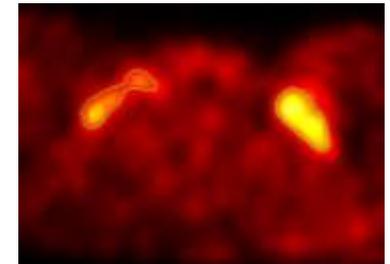
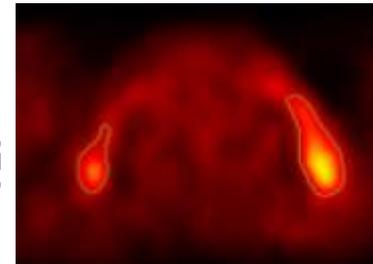
TDM

## Discrimination

TDM

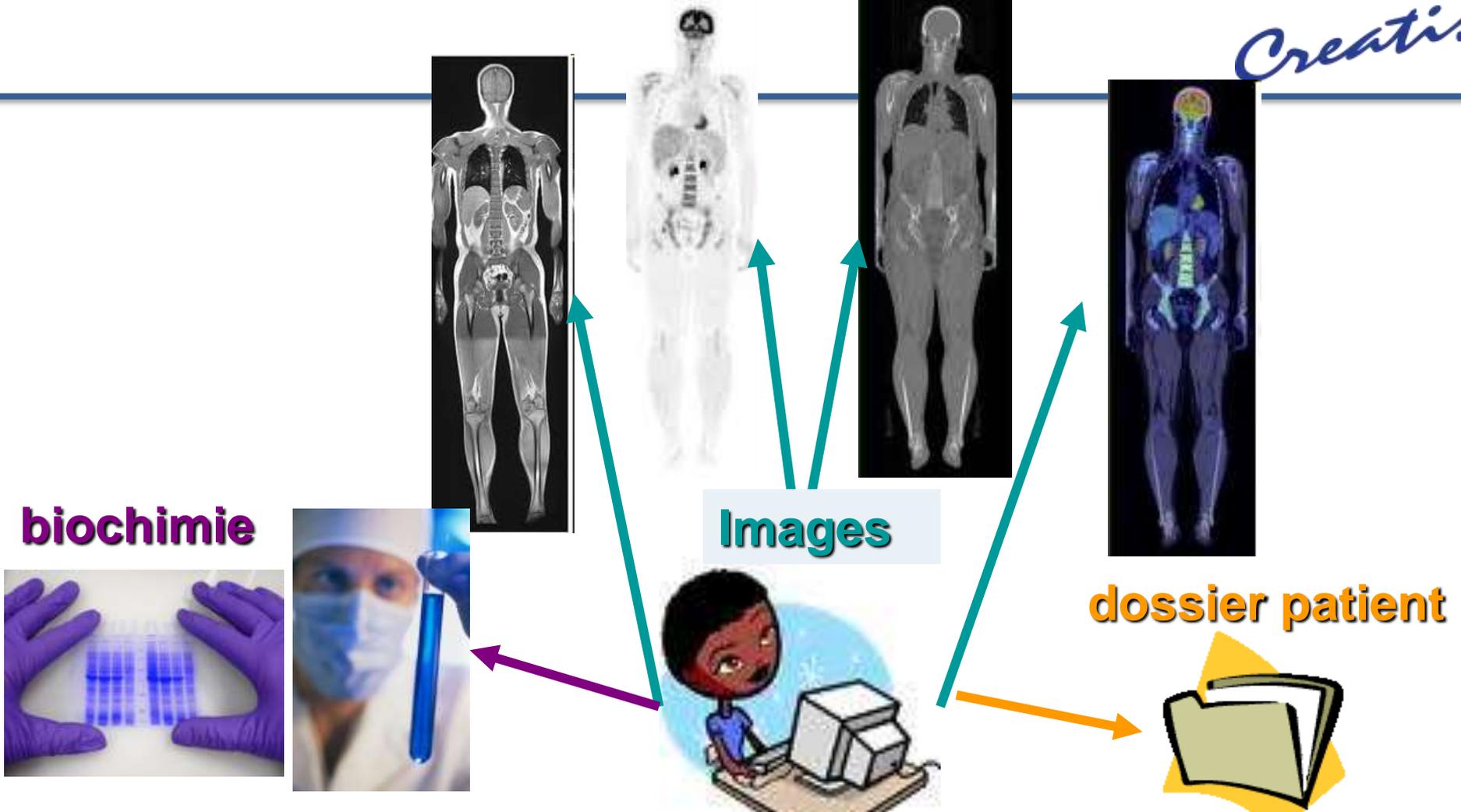


TEP



normal

pathologique



biochimie

Images

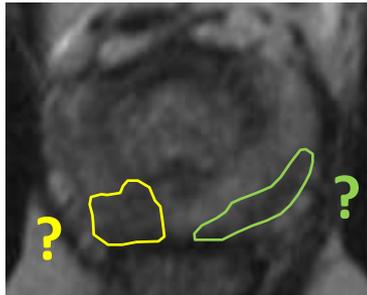
dossier patient

Traitement d'images, extraction de métriques, prise de décision

APPRENDRE à partir de cas connus

- Place de l'imagerie médicale dans le diagnostic du cancer
  - de plus en plus de données...
- **Imagerie médicale et apprentissage pour l'aide au diagnostic**
  - Principe des systèmes d'aide au diagnostic
  - Application en imagerie IRM du cancer de la prostate
- Quelques enjeux de l'apprentissage pour l'aide au diagnostic

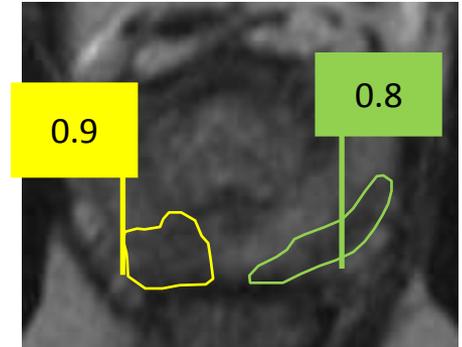
## ■ Les systèmes d'aide au **diagnostic**



1. Les ROI suspectes sont contourées sur l'image

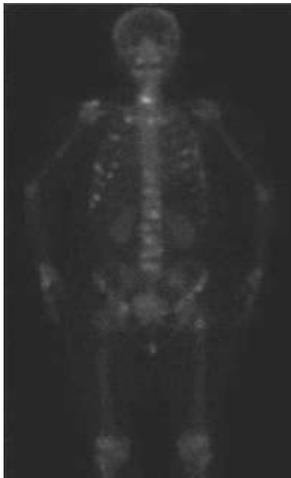


CADx



3. Le CAD retourne un score de malignité pour chaque ROI

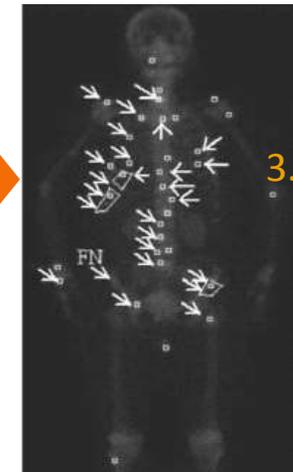
## ■ Les systèmes d'aide à la **détection**



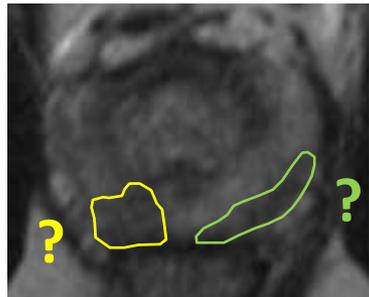
1. Le ou les images sont entrées sans annotation



CADe



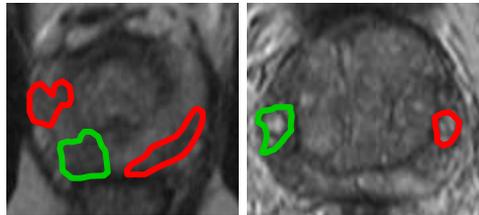
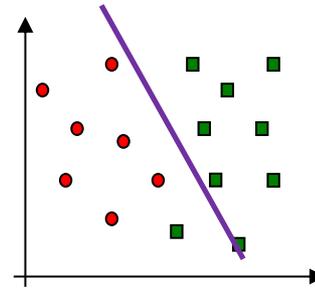
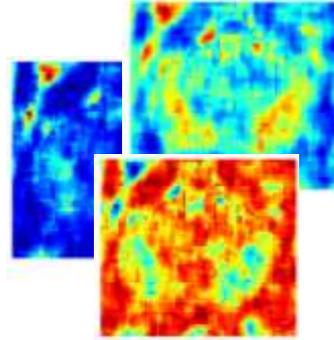
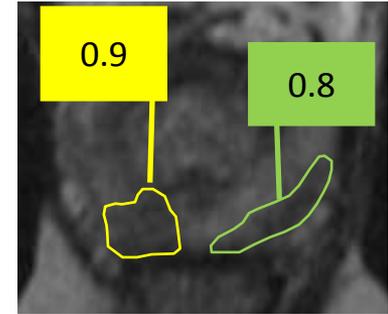
3. Le CAD renvoie une cartographie des zones pathologiques



Extraction de caractéristiques

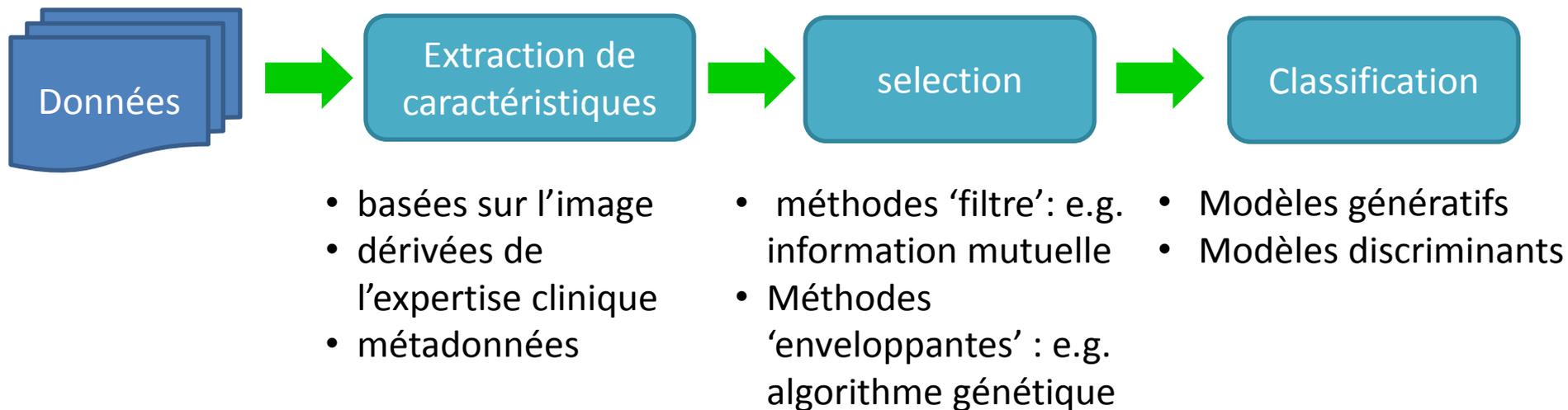


Classification

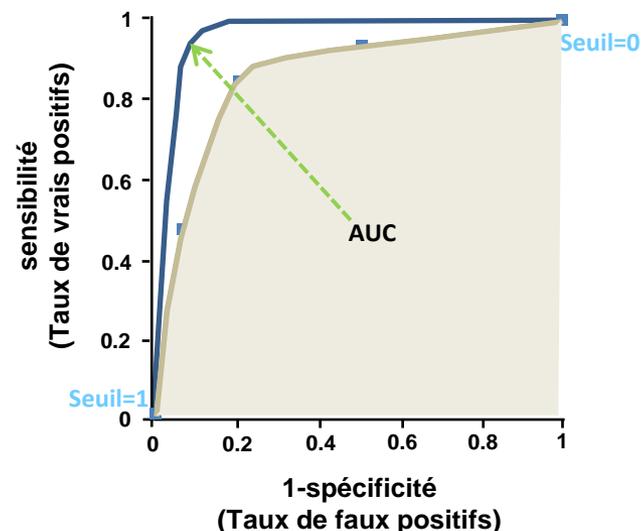


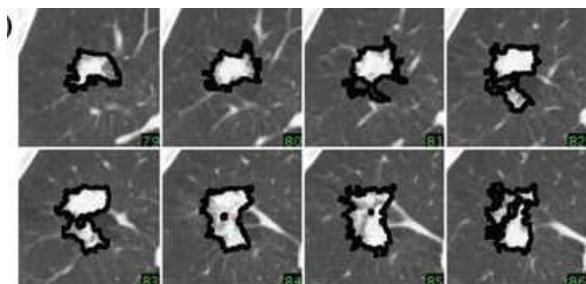
Apprentissage



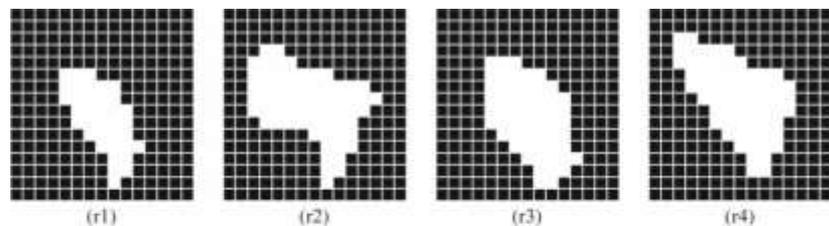


- Optimisation des différentes étapes basée sur une métrique dérivées de l'analyse psychophysique (courbe ROC, sensibilité, spécificité)





Contourage manuel de lésions pulmonaires en imagerie TDM



Contourage manuel d'une lésion réalisé par 4 experts

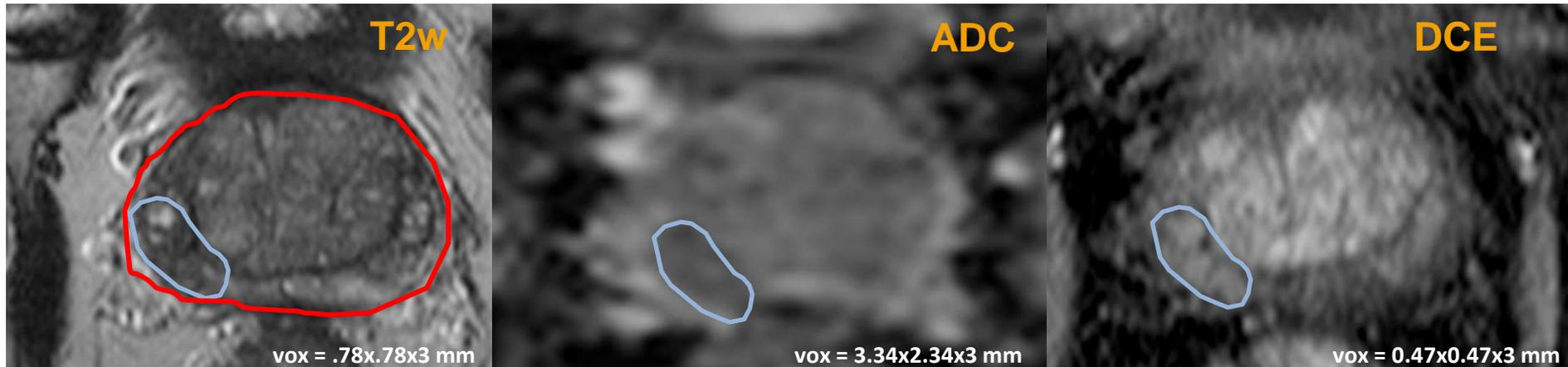
[Biancardi, IJCARS 2010]

- Données hétérogènes, corrélées, bruitées
- Annotations difficilement accessibles: quelle vérité terrain?

- Place de l'imagerie médicale dans le diagnostic du cancer
  - de plus en plus de données...
- **Imagerie médicale et apprentissage pour l'aide au diagnostic**
  - Principe des systèmes d'aide au diagnostic
  - Application en imagerie IRM du cancer de la prostate
- Quelques enjeux de l'apprentissage pour l'aide au diagnostic

## Contexte

- Gain de l'IRM multi-séquences démontré pour le diagnostic *in vivo*
- Analyse des données complexe
- grande variabilité inter- et intra-lecteur



*Hyposignal T2 ?*

*Restriction de la diffusion?*

*Hyper-vascularité ?*

**Objectif :** Discriminer les lésions cancéreuses des lésions bénignes mais suspectes (adénome, inflammation..)



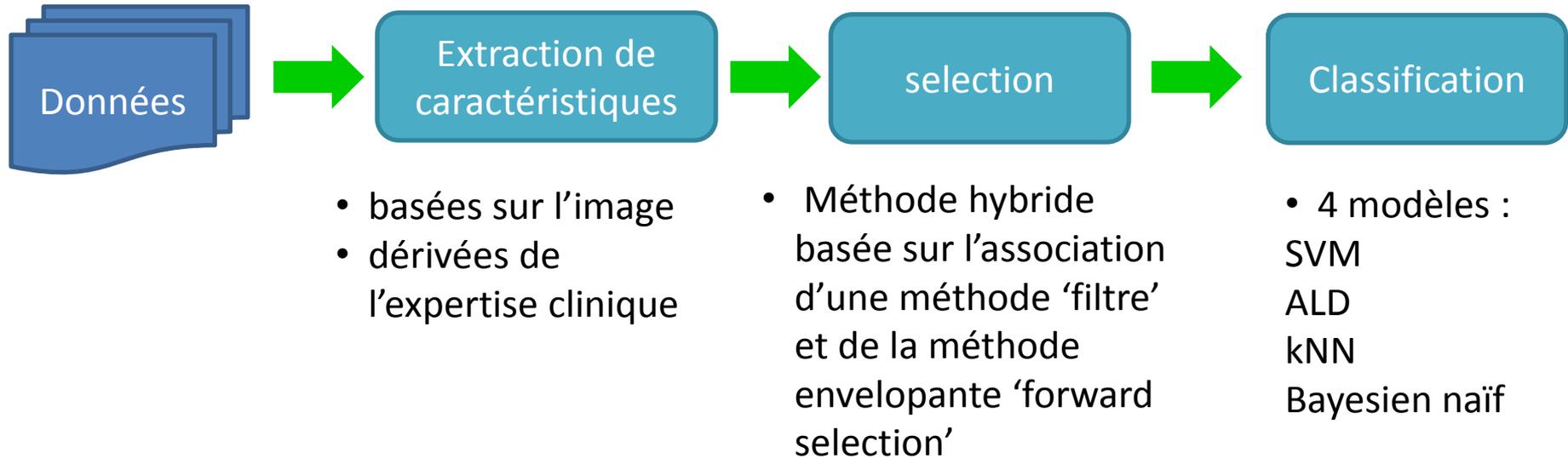
**Coupe histologique**

Collaboration Pr Olivier Rouvière, HCL, LabTau  
Projets INCa: Cartographix (2011-14) et LYRIC (2012-17)  
Base de données CLARA-P

## Méthode

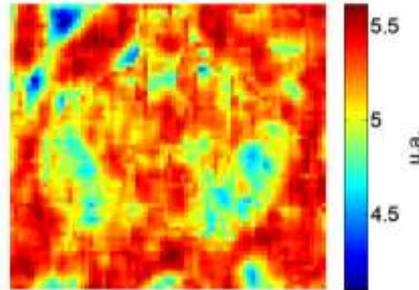
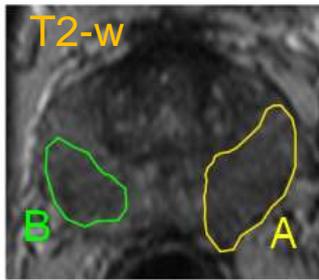
- Extraction et sélection de caractéristiques discriminantes
- Analyse exhaustive de différents schémas de classification (algorithme de classification + série de caractéristiques)

**Verrous :** Apprentissage sur un faible nombre d'échantillons



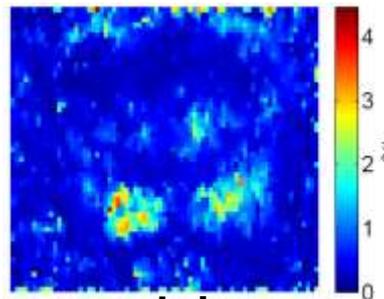
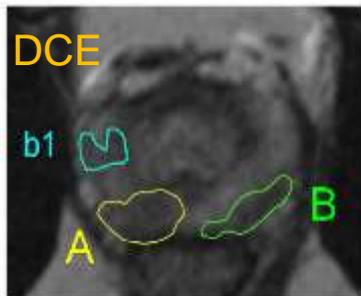
## Extraction de caractéristiques

### Paramètres de texture



entropie

### Paramètres cinétiques semi-quantitatifs



wash-in

- 42 régions d'intérêt (ROI) cancer et 49 ROI bénignes mais suspectes
- Extraction de 117 caractéristiques : statistique 1<sup>er</sup> et 2<sup>nd</sup> ordre, gradient, paramètres pharmacocinétiques

[Niaf, PMB 2012]

## Performances intrinsèques

	AUC	IC
SVM	0.72	[0.61-0.82]
LDA	0.56	[0.44-0.67]
KNN	0.66	[0.54-0.77]
NBC	0.63	[0.51-0.74]

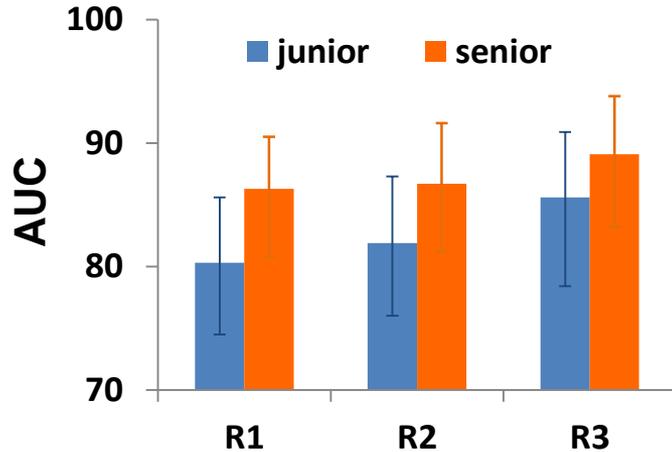
	# caract.	AUC	IC
SVM	15	0.82	[0.73-0.90]
LDA	15	0.75	[0.64-0.84]
KNN	12	0.78	[0.68-0.87]
NBC	4	0.77	[0.66-0.85]

Aire sous la courbe ROC

[Niaf, PMB 2012]

- Performances intrinsèques > état de l'art
- Apport des différentes séquences IRM
- Apport de la sélection d'attributs (dans le contexte de l'apprentissage sur un faible nombre de cas)

## Impact sur les performances diagnostiques



- Etude psychophysique sur 12 radiologues
- R1 et R2 : lectures sans CAD
- R3 : lecture avec CAD

[Niaf, Radiology 2014]

Collaboration Service Biostatistiques, HCL, Lyon

## Conclusion

- Bonne performance intrinsèque et diagnostique
- Plateau de performance intrinsèque (AUC~0.8)

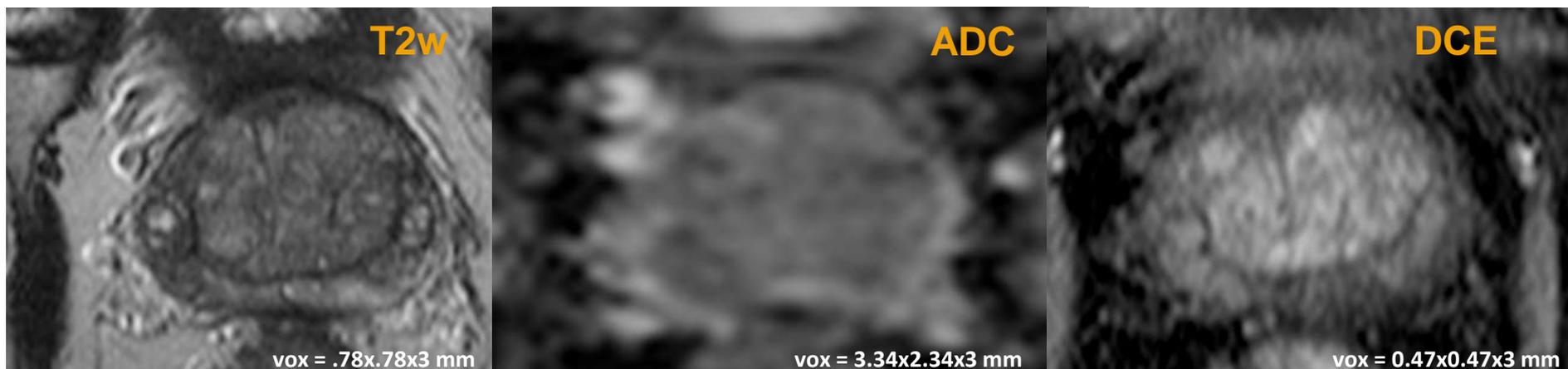
→ Aller au delà des schémas classiques de classification

**Objectif :** Fournir une carte probabiliste de la pathologie dans l'image

## Méthode

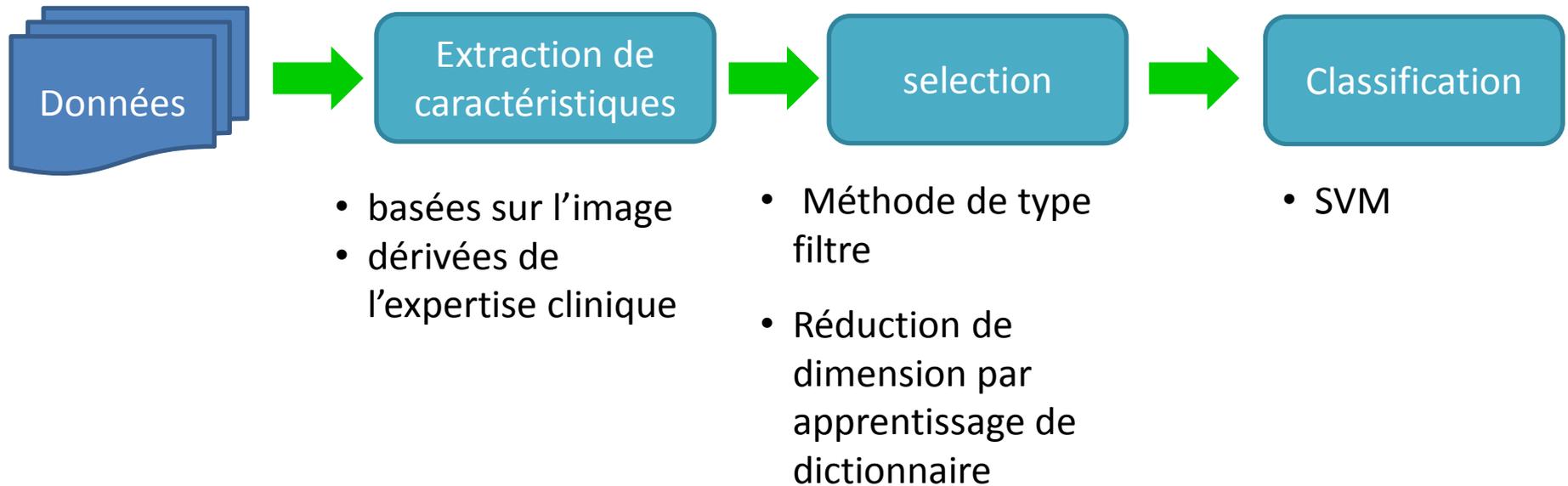
- Extraction et sélection de caractéristiques discriminantes
- Analyse exhaustive de différents schémas de classification

**Verrous :** Apprentissage sur des masses de données (voxels) déséquilibrées



Collaboration Pr Olivier Rouvière, HCL, LabTau  
Projets INCa: Cartographix (2011-14) et LYRIC (2012-17)  
Base de données CLARA-P

## Choix et contributions méthodologiques



## Données

- 42 régions d'intérêt (ROI) cancer, 49 ROI bénignes mais suspectes (NS) et 124 ROIs normales (N)

	N	NS	GS6	GS7	GS8	GS9
voxels (x 10 <sup>e3</sup> )	360	32	11	34	16	7

## Extraction de caractéristiques

- Méthode d'apprentissage de dictionnaire non supervisé et semi-supervisé (Mairal et al, PAMI 2012)

$$\hat{\mathbf{x}}_i = \sum_{k=1}^K \alpha_i^k \mathbf{d}_k = \mathbf{D}\boldsymbol{\alpha}_i$$

$$\boldsymbol{\alpha}_i \in \mathbb{R}^K$$

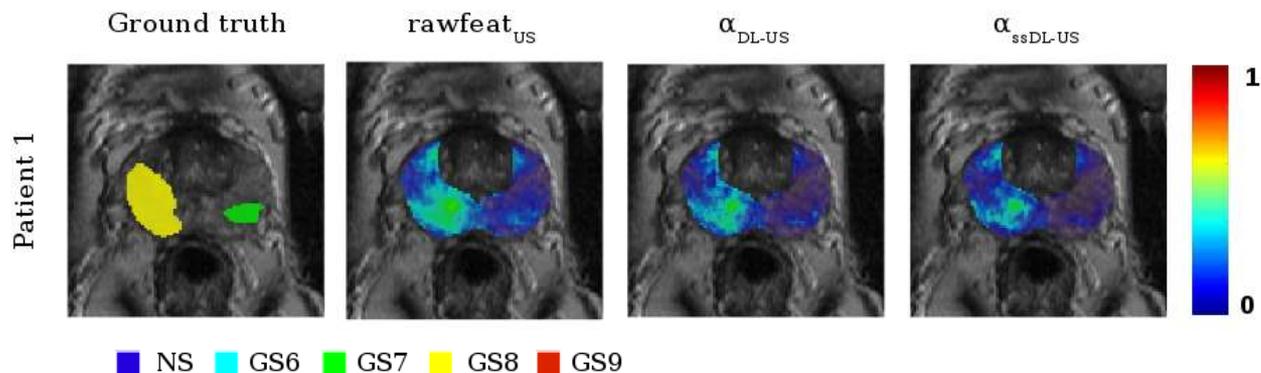
$$\hat{\mathbf{x}}_i \in \mathbb{R}^P$$

$$\min_{\mathbf{D} \in \mathbb{R}^{p \times k}, \boldsymbol{\alpha}_i \in \mathbb{R}^K} \frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}_i\|_1$$

$$\mathbf{x}_i \in \mathbb{R}^P$$

## Résultats

Cartographie du cancer de prostate avec un SVM linéaire



[Lehaire et al, ICIP 2014]  
(Collab OCA Lagrange Nice)

## Conclusion

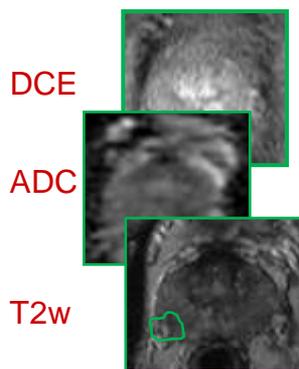
- Bonnes performances intrinsèques
- Prédominance de l'influence non linéarité sur le déséquilibre des classes
- Pas d'amélioration significative de performances avec l'apprentissage de dictionnaire mais réduction de dimension → amélioration des performances calculatoires

**Perspectives:** Caractérisation in vivo de la pathologie (score d'agressivité..)

- Place de l'imagerie médicale dans le diagnostic du cancer
  - de plus en plus de données...
- Imagerie médicale et apprentissage pour l'aide au diagnostic
  - Principe des systèmes d'aide au diagnostic
  - Application en imagerie IRM du cancer de la prostate
- **Quelques enjeux de l'apprentissage pour l'aide au diagnostic**

## Position du problème

- Apprentissage sur des bases de données issues de différents systèmes d'acquisition ou différents protocoles d'imagerie



Données source  $D_s$

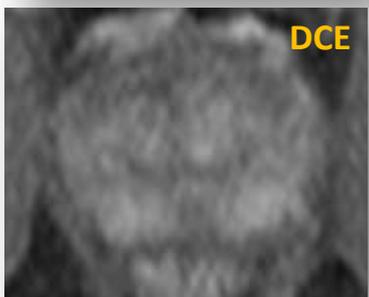
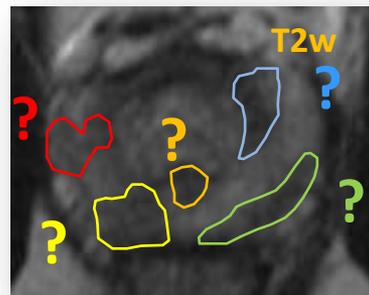


Données cible  $D_t$

## Objectifs

- $D_s$  et  $D_t$  ont un nombre suffisant d'exemples d'apprentissage :
  - Améliorer les performances diagnostiques en fusionnant les bases
- $D_t$  contient peu d'exemples d'apprentissage :
  - Adapter le modèle appris sur  $D_s$

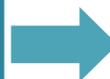
## Approche classique



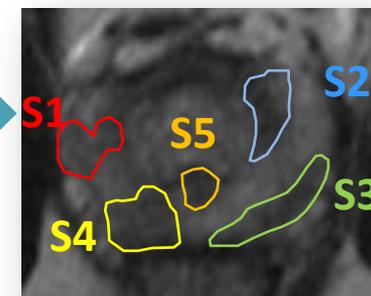
Extraction de caractéristiques des différents types d'images

+

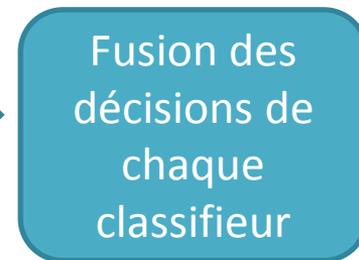
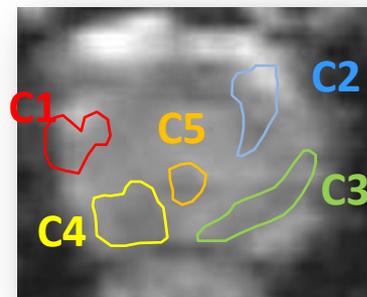
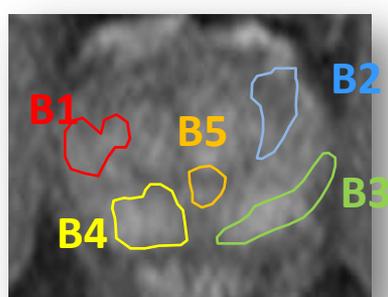
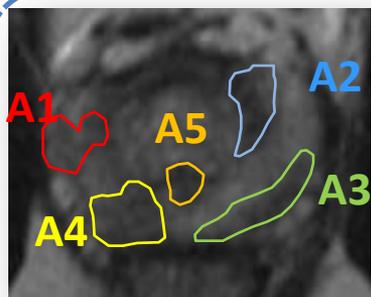
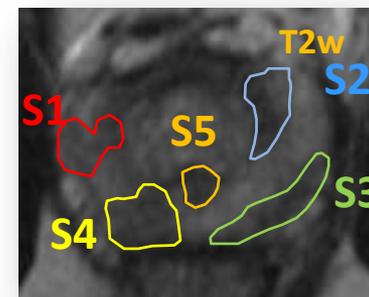
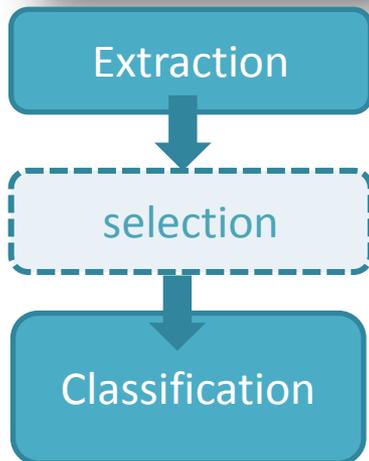
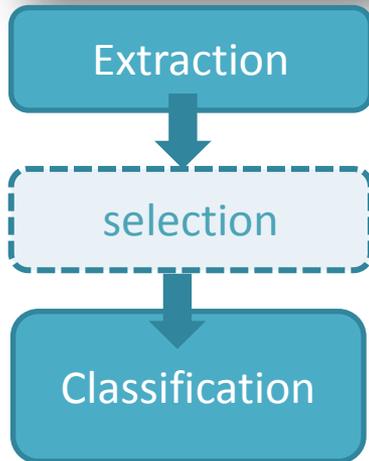
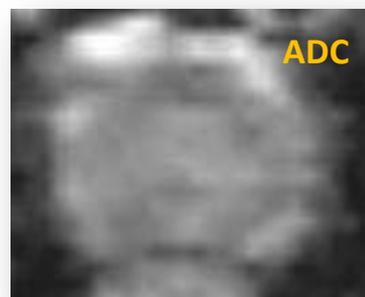
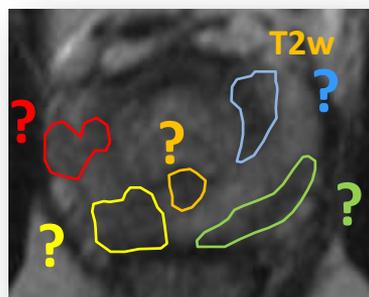
selection



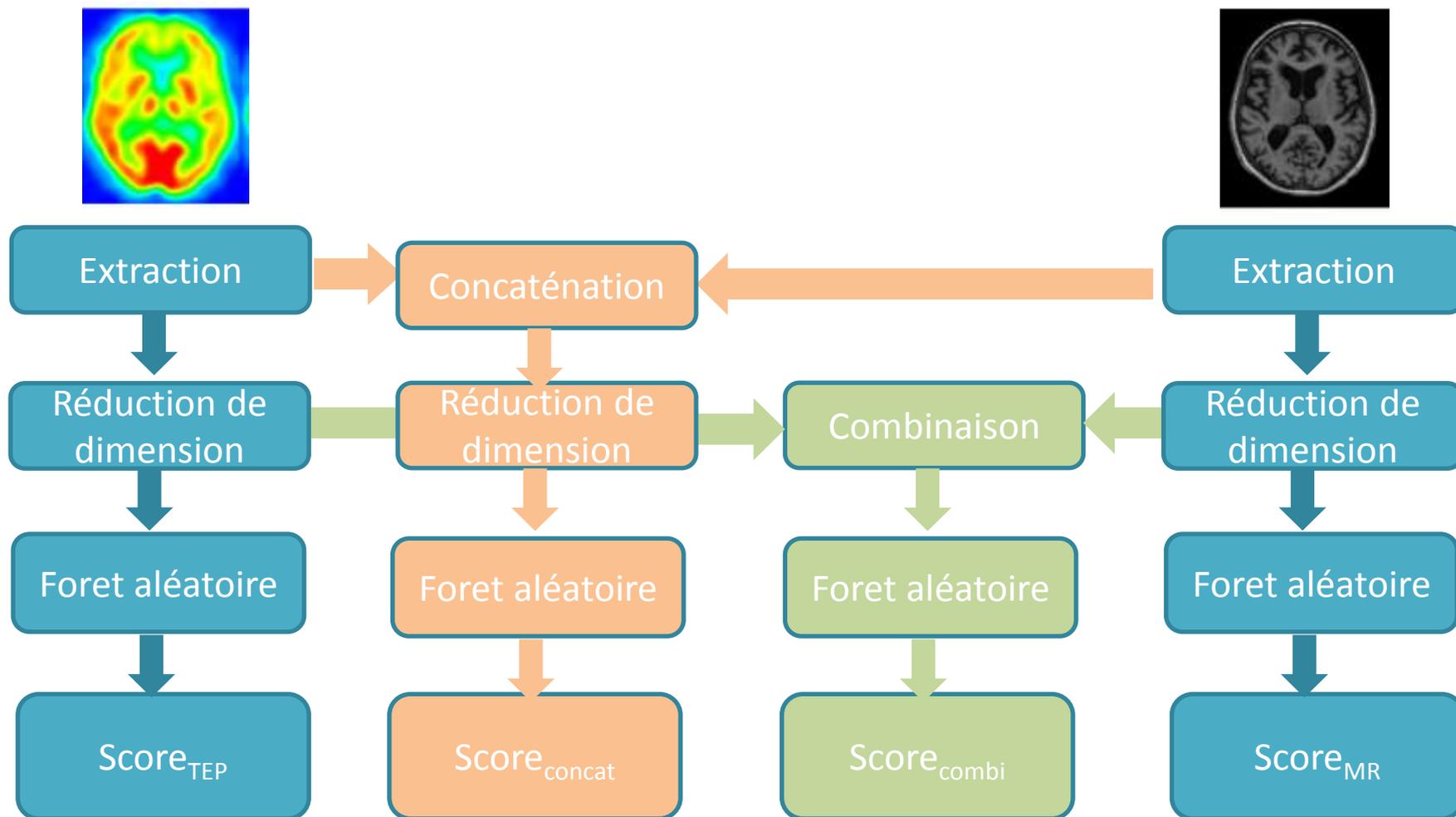
Classification



## Fusion de décision



## Fusion de caractéristiques



[K. Gray, PhD thesis, UCL, London, 2012]

	AD vs. HC		MCI vs. HC	
	MRI	FDG-PET	MRI	FDG-PET
Acc. (%)	87.2 (2.0)	87.8 (2.6)	64.8 (3.0)	65.3 (1.9)
Bacc. (%)	87.2 (2.9)	87.8 (4.0)	65.2 (5.3)	65.3 (3.4)
Sens. (%)	87.5 (3.2)	91.8 (2.9)	64.8 (3.9)	65.3 (2.9)
Spec. (%)	86.9 (2.6)	83.8 (5.1)	65.5 (6.7)	65.2 (3.8)

◀ Avec réduction de dimension

	AD vs. HC		MCI vs. HC	
	Joint embedding	Concatenation	Joint embedding	Concatenation
Acc. (%)	90.0 (2.6)	87.9 (2.6)	75.5 (2.2)	64.3 (2.4)
Bacc. (%)	89.4 (3.6)	87.9 (4.3)	74.7 (3.0)	66.9 (3.9)
Sens. (%)	88.9 (3.4)	92.0 (3.9)	76.9 (3.2)	59.8 (3.2)
Spec. (%)	89.8 (3.8)	83.8 (4.7)	72.4 (4.5)	74.0 (4.6)

◀ Fusion des modalités:  
concaténation avant  
réduction de dimension  
(Score<sub>concat</sub>) ou après  
(Score<sub>combi</sub>)

AD = Alzheimer disease

HC = Healthy control

MCI =mild cognitive impairment

[K. Gray, PhD thesis, UCL, London, 2012]

- Explosion des temps de calcul → Accès à des ressources de calcul distribué

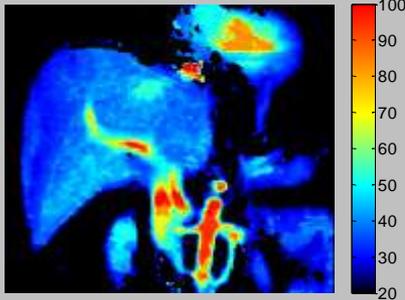
## Virtual Imaging Platform

<http://www.creatis.insa-lyon.fr/vip>

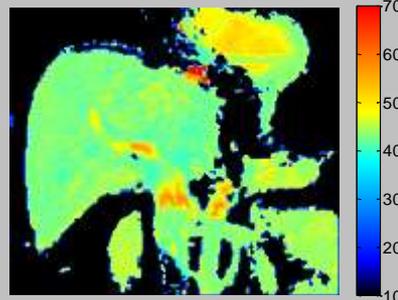


### Exemple d'application

Débit artériel  
( $\text{mL} \cdot \text{min}^{-1} \cdot 100\text{g}^{-1}$ )



Index de perfusion  
hépatique (%)

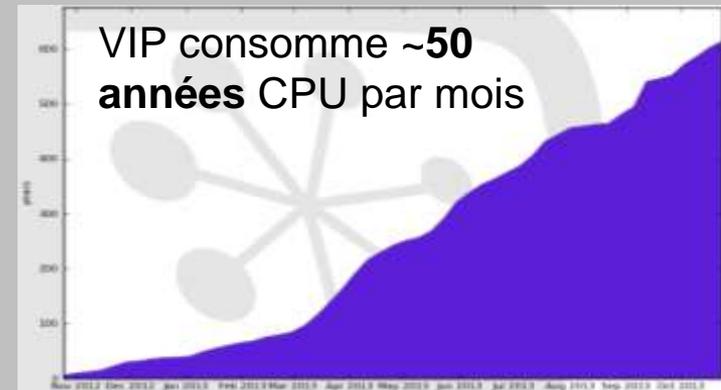


Cartographies des paramètres de perfusion hépatique  
1 an de calcul réalisé en 9.5 jours sur la grille EGI

[Leporq, J Med Eng 2013]

### Infrastructure

Supported by EGI Infrastructure  
Uses biomed VO (most used EGI VO for life sciences in 2013)  
VIP accounts for ~25% of biomed's activity



VIP consomme ~50  
années CPU par mois

France-Grilles



DIRAC

### Utilisateurs



479 utilisateurs en Nov 2013  
(175 en France)

- Place croissante de l'imagerie multimodale dans le diagnostic du cancer
- Apport des méthodes par apprentissage pour l'aide au diagnostic en imagerie du cancer
- Limites des systèmes d'aide au diagnostic basés sur un schéma classique
- Quelques pistes pour adapter les méthodes par apprentissage aux spécificités de l'imagerie médicale