Mixture of multivariate multiple-scaled Student distributions : application to the characterization of brain tumors.

Alexis Arnaud

Supervisors : F. Forbes, E. Barbier, B. Lemasson
Laboratories : LJK, INRIA & GIN
Funding : UJF

1st year-PhD
30 June 2015

Outline

Outline

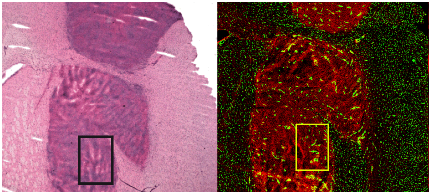
How to characterize brain tumor ?

Healthy ? ◀

▶ Not healthy ?

How to characterize brain tumor ?

Histology vs multiparametric MRI

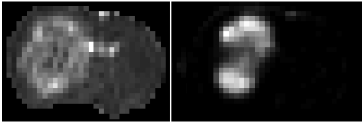


Histology section

- \+ gold standard : provides the ground truth
- \+ precise information
- – local information : only parts of the tumor may be sampled
- – invasive operation (biopsy), not always feasible

How to characterize brain tumor ?

Histology vs multiparametric MRI



2 MRI maps

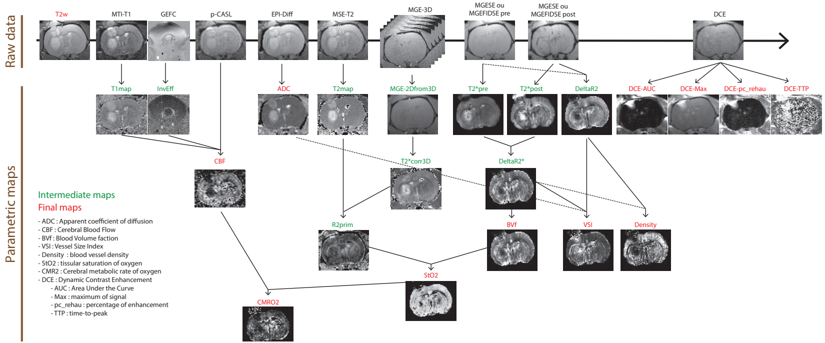
- − under development method
- + global information : whole tumor visible
- + non-invasive operation

Goal : find the voxels inside the MRI maps which belong to the tumor, in order to characterize a tumor, to avoid invasive biopsies.

How to characterize brain tumor ?

Problem of multiparametric MRI

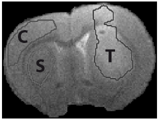
How to extract information from all of the parametric maps ?



▶ **Approach** : multivariate clustering with mixture models.

How to characterize brain tumor ?

Previous study (Coquery *et al.* 2014)

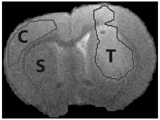


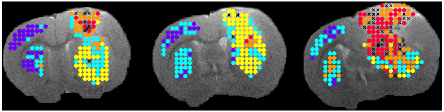
3 Regions of interest were manually delimited on each MRI map :

- 2 healthy regions on the left (Cortex and Striatum)
- 1 tumor region on the right (Tumor)

How to characterize brain tumor ?

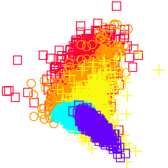
Previous study (Coquery *et al.* 2014)





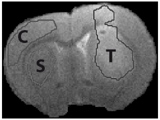
Voxels were partitioned using a Gaussian mixture into a number of classes :

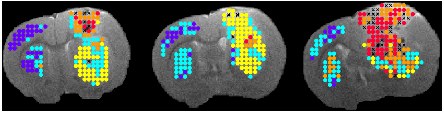
- + link between classes and tissue types
- − tissue values may not have Gaussian shapes



How to characterize brain tumor ?

Previous study (Coquery *et al.* 2014)

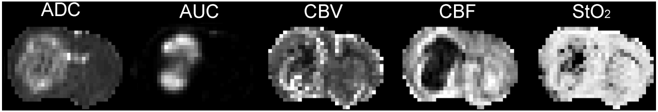




- Proposed approach : classification of voxels using a mixture of heavy-tailed distributions, the multivariate multiple-scaled Student distributions (MMSD).

Outline

Multiparametric MRI data



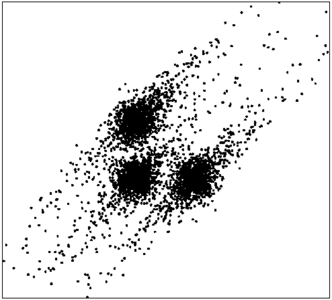
- 5 physiological parameters :
  - **ADC** : apparent diffusion coefficient
  - **CBV** : cerebral blood volume
  - **CBF** : cerebral blood flow
  - **AUC** : blood vessel permeability
  - **StO_2** : oxygen saturation

- 5 dimensional data set :

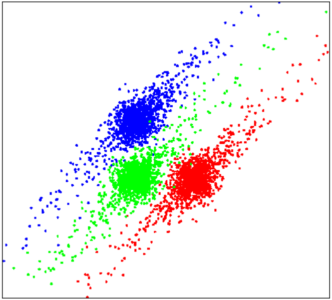
  $\boldsymbol{Y} = \{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N\}$ the set of all voxels, of size N
  with $\boldsymbol{Y}_n = \{\boldsymbol{Y}_{n,\text{ADC}}, \ldots, \boldsymbol{Y}_{n,\text{StO}_2}\}$ the mesures on the nth voxel

Clustering of MRI data

We supose that data arise from K different classes (for K different tissues), and we want to recover those classes.



Simulated data



Latent classification

Clustering of MRI data via mixture modeling

Let $\boldsymbol{Z}$ be the latent variable which links one observation to one class :

$$\begin{cases} (\boldsymbol{Y}_n \,|\, \boldsymbol{Z}_n = k) & \sim \quad f_k(\boldsymbol{\theta}_k) \\ \boldsymbol{Z}_n & \sim \quad \mathcal{M}(\pi_1, \ldots, \pi_K) \end{cases} \tag{1}$$

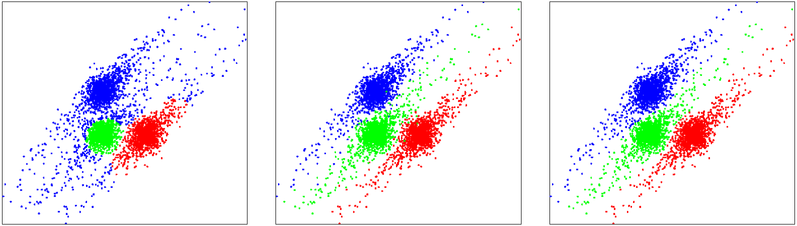
$\Leftrightarrow$

$$\Pr(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k f_k(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}_k) \tag{2}$$

Clustering of MRI data via mixture modeling

In the previous study, $f_k(\boldsymbol{\theta}_k) = \mathcal{N}_5(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with $\boldsymbol{\mu}_k \in \mathbb{R}^5$ and $\boldsymbol{\Sigma}_k \in \mathcal{S}^+_{5\times 5}(\mathbb{R})$: lack of flexibility in cluster shape modeling (Coquery *et al.* - 2014).

In this study, $f_k(\boldsymbol{\theta}_k)$ is a heavy-tailed distribution : a multivariate multiple-scaled Student distribution (MMSD).



Gaussian Simulation MMSD

Outline

Clustering

Gaussian mixture

- Mixture model :

$$\Pr(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k f_k(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}_k) \quad (3)$$

$$f_k(\boldsymbol{\theta}_k) \sim \mathcal{N}_M(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (4)$$

With : $\boldsymbol{\mu}_k \in \mathbb{R}^M, \boldsymbol{\Sigma}_k \in \mathcal{S}^+_{M\times M}(\mathbb{R})$

- Decomposition of the covariance matrix for a better flexibility :

$$\boldsymbol{\Sigma}_k = \boldsymbol{U}_k \boldsymbol{D}_k \boldsymbol{U}_k^{t} \quad (5)$$

with :

$\boldsymbol{U}_k \in \mathcal{O}(M)$ the orthogonal matrix of eigenvectors
$\boldsymbol{D}_k \in \mathcal{D}(M)$ the diagonal matrix of eigenvalues

Clustering

Gaussian mixture

- Decomposition of the covariance matrix :

$$\Pr(\boldsymbol{y}_n ; \boldsymbol{\theta}) = \sum_{k=1}^{\mathrm{K}} \pi_k f_k(\boldsymbol{y}_n ; \boldsymbol{\theta}_k) \quad (3)$$

$$f_k(\boldsymbol{\theta}_k) \sim \mathcal{N}_{\mathrm{M}}(\boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k) \quad (4)$$

With : $\boldsymbol{\mu}_k \in \mathbb{R}^{\mathrm{M}}, \boldsymbol{U}_k \in \mathcal{O}(\mathrm{M}), \boldsymbol{D}_k \in \mathcal{D}(\mathrm{M})$

- Tractable estimation using an Expectation-Maximization (EM) algorithm and a minimization algorithm of type Flury and Gautschi for the orthogonal matrices.
- The Gaussian tail is very short and sensitive to atypical observations (outliers).

Clustering

Student mixture

Mixture model with one-dimensional Student distributions :

$$\begin{aligned} \Pr\left(y_n \; ; \; \boldsymbol{\theta}\right) &= \sum_{k=1}^{K} \pi_k f_k\left(y_n \; ; \; \boldsymbol{\theta}_k\right) \\ f_k\left(\boldsymbol{\theta}_k\right) &\sim t_1\left(\mu_k, \sigma_k^2, \nu_k\right) \\ t_1\left(y_n \; ; \; \mu_k, \sigma_k^2, \nu_k\right) &= \frac{\Gamma\left(\frac{\nu_k+1}{2}\right)}{\sigma_k \Gamma\left(\frac{\nu_k}{2}\right)\left(\pi\nu_k\right)^{\frac{1}{2}}} \left[1 + \frac{\left(y_n - \mu_k\right)^2}{\nu_k \sigma_k^2}\right]^{-\frac{\nu_k+1}{2}} \end{aligned} \quad (5)$$

- How to get a multidimensional Student distribution ?

Standard multivariate Student distribution

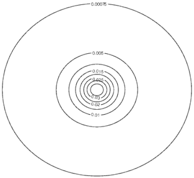
One possible generalization to a M-dimensional distribution

$$p_{\mathrm{MS}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu\right) = \frac{\Gamma\left(\frac{\nu+\mathrm{M}}{2}\right)}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}\,\Gamma\left(\frac{\nu}{2}\right)\left(\pi\nu\right)^{\frac{\mathrm{M}}{2}}}\left[1+\frac{\delta^2\left(\boldsymbol{y},\boldsymbol{\mu},\boldsymbol{\Sigma}\right)}{\nu}\right]^{-\frac{\nu+\mathrm{M}}{2}}$$

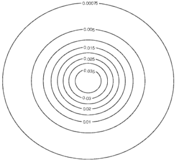
with : $\delta^2\left(\boldsymbol{y},\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = \left(\boldsymbol{y}-\boldsymbol{\mu}\right)^{\mathrm{t}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}\right)$

the Mahalanobis distance

- But the degree of freedom remains scalar.



$\nu = 5$



$\nu = 10$

Standard multivariate Student distribution

Useful representation : infinite mixture of scaled Gaussians

$$p_{\mathrm{MS}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu\right) = \frac{\Gamma\left(\frac{\nu+\mathrm{M}}{2}\right)}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}\,\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{\frac{\mathrm{M}}{2}}}\left[1+\frac{\delta^2\left(\boldsymbol{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)}{\nu}\right]^{-\frac{\nu+\mathrm{M}}{2}}$$

$$= \int_0^\infty \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \frac{1}{w}\boldsymbol{\Sigma}\right)\mathcal{G}\left(w\ ;\ \frac{\nu}{2}, \frac{\nu}{2}\right)\mathrm{d}w$$

with :

- $\mathcal{N}_{\mathrm{M}}$ the M-multivariate Gaussian distribution
- $\mathcal{G}$ the Gamma distribution
- the real latent variable W is called the weight

Multivariate multiple-scaled Student distribution

Use of the eigenvalue decomposition to get a multidimensional degree of freedom (Forbes and Wraith - 2014)

Let U and D the eigenvalue decomposition of the symmetric positive definite $\mathrm{T} = \boldsymbol{\Sigma}^{-1}$: $\mathrm{T} = \mathrm{UDU}^{\mathrm{t}}$.

$$p_{\mathrm{MS}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{U}, \boldsymbol{D}, \nu\right) = \int_{\mathbb{R}_+^*} \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \frac{1}{w}\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{t}}\right) \mathcal{G}\left(w\ ;\ \frac{\nu}{2}, \frac{\nu}{2}\right)\, \mathrm{d}w$$

Multivariate multiple-scaled Student distribution

Use of the eigenvalue decomposition to get a multidimensional degree of freedom (Forbes and Wraith - 2014)

Let U and D the eigenvalue decomposition of the symmetric positive definite $\mathrm{T} = \boldsymbol{\Sigma}^{-1}$: $\mathrm{T} = \mathrm{UDU}^{\mathrm{t}}$.
Let $\boldsymbol{W} \in \mathbb{R}_{+*}^{\mathrm{M}}$ a M-dimensional weight (one scalar weight for each dimension) :

$$p_{\mathrm{MMS}}\left(\boldsymbol{y} \; ; \boldsymbol{\mu}, \boldsymbol{U}, \boldsymbol{D}, \boldsymbol{\nu}\right) =$$

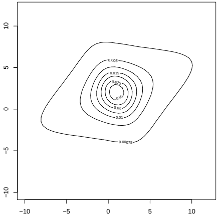
$$\int_{\mathbb{R}_{+*}^{\mathrm{M}}} \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y} \; ; \boldsymbol{\mu}, \boldsymbol{U} \mathrm{diag}\left(\mathbf{w}\right)^{-1} \boldsymbol{D}^{-1} \boldsymbol{U}^{\mathrm{t}}\right) \prod_{m=1}^{\mathrm{M}} \mathcal{G}\left(w_m \; ; \frac{\nu_m}{2}, \frac{\nu_m}{2}\right) \, \mathrm{d}\mathbf{w}$$

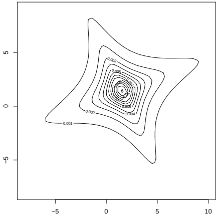
Multivariate multiple-scaled Student distribution

Use of the eigenvalue decomposition to get a multidimensional degree of freedom (Forbes and Wraith - 2014)

$$p_{\mathrm{MMS}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{U}, \boldsymbol{D}, \boldsymbol{\nu}\right) =$$

$$\int_{\mathbb{R}_{+*}^{\mathrm{M}}} \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{U}\mathrm{diag}\left(\mathbf{w}\right)^{-1}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{t}}\right) \prod_{m=1}^{\mathrm{M}} \mathcal{G}\left(w_m\ ;\ \frac{\nu_m}{2}, \frac{\nu_m}{2}\right)\ \mathrm{d}\mathbf{w}$$





Outline

Estimation of a MMSD mixture

Mixture of MMSD

- density of the model :

$$p\left(\boldsymbol{y}_n\ ;\ \boldsymbol{\theta}\right) = \sum_{k=1}^{\mathrm{K}} \pi_k p_{\mathrm{MMSD}}\left(\boldsymbol{y}_n\ ;\ \boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k, \boldsymbol{\nu}_k\right)$$

with :

$$\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{\mathrm{K}}\}$$

$$\boldsymbol{\theta}_k = (\boldsymbol{\pi}_k, \boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k, \boldsymbol{\nu}_k)$$

- estimation of parameters :
  - a simple EM algorithm
  - with a Flury and Gautschi algorithm for the orthogonal matrices

Estimation of a MMSD mixture

EM algorithm

For $k \in \{1, \ldots, \mathrm{K}\}$, and $n = \{1, \ldots, \mathrm{N}\}$:

- Z_n is the unknown class variable : $\Pr(\mathrm{Z}_n = k) = \pi_k$
- $\boldsymbol{W}_n$ is the unknown weight
- $\boldsymbol{\Delta}_n = \mathrm{diag}(\boldsymbol{W}_n) = \mathrm{diag}(W_{n1}, \ldots, W_{n\mathrm{M}}) \in \mathbb{R}_{+*}^{\mathrm{M}}$

$$\begin{aligned}
\boldsymbol{Y}_n \,|\, \mathrm{Z}_n = k, \boldsymbol{W}_n = \mathbf{w}_n \,;\, \boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k &\sim \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{\mu}_k, \boldsymbol{U}_k \boldsymbol{\Delta}_n{}^{-1} \boldsymbol{D}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}}\right) \\
\boldsymbol{W}_n \,|\, \mathrm{Z}_n = k \,;\, \boldsymbol{\nu}_k &\sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G}\left(\frac{\nu_{km}}{2}, \frac{\nu_{km}}{2}\right) \\
\mathrm{Z}_n \,;\, \boldsymbol{\pi} &\sim \mathcal{M}\left(1, \pi_1, \ldots, \pi_{\mathrm{K}}\right)
\end{aligned}$$

Estimation of a MMSD mixture

Outliers accommodation

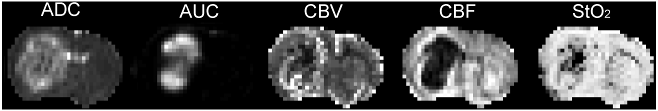
At the E-step of the iteration r, the conditional density use :

$$\begin{aligned} \mathbf{w}_{nk}^{(r)} &= \left(w_{nk1}^{(r)}, \ldots, w_{nk\mathrm{M}}^{(r)}\right) \\ \text{with : } w_{nkm}^{(r)} &= \mathbb{E}\left[W_{nm} \,|\, Z_n = k, \boldsymbol{y}_n \,;\, \boldsymbol{\theta}^{(r)}\right] \\ &= \frac{\nu_{km}^{(r)} + 1}{\nu_{km}^{(r)} + \left[\boldsymbol{D}_k^{(r)}\right]_m \left[\boldsymbol{U}_k^{(r)^{\mathrm{t}}} \left(\boldsymbol{y}_n - \boldsymbol{\mu}_k^{(r)}\right)\right]_m^2} \end{aligned} \tag{6}$$

The higher $\left[\boldsymbol{D}_k^{(r)}\right]_m \left[\boldsymbol{U}_k^{(r)^{\mathrm{t}}} \left(\boldsymbol{y}_n - \boldsymbol{\mu}_k^{(r)}\right)\right]_m^2$, the more the observation is atypical.

Outline

Multiparametric MRI data

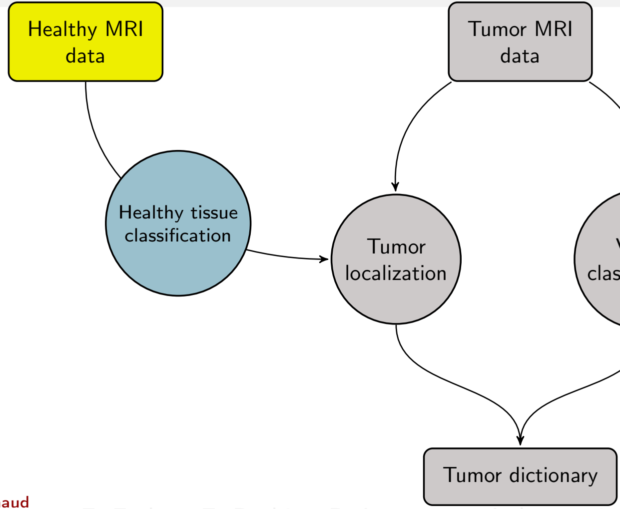


- data in **dimension 5**
- **26 parameters per class**
- estimation by **EM** and **Flury & Gautschi** or **Stiefel manifold optimization** algorithms
- choice of the number of classes according with **BIC** and **ICL** criterions

- **8 healthy rats** (49 000 voxels)
- **37 rats with tumors** (290 000 voxels)
- 4 tumor models : **9L**, **C6**, **F98**, **RG2**

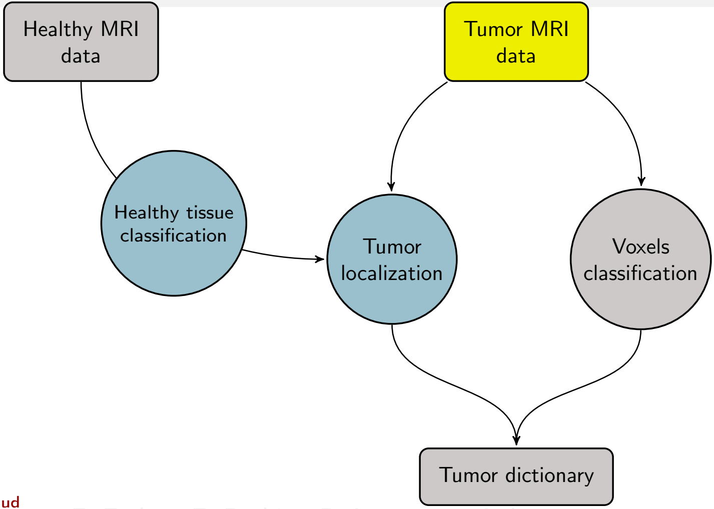
Processing pipeline

Healthy voxels classification



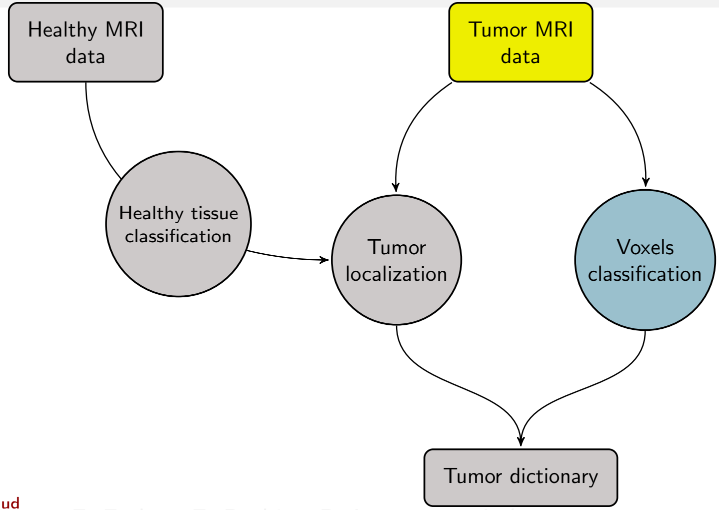
Processing pipeline

Tumor localization (ROI)



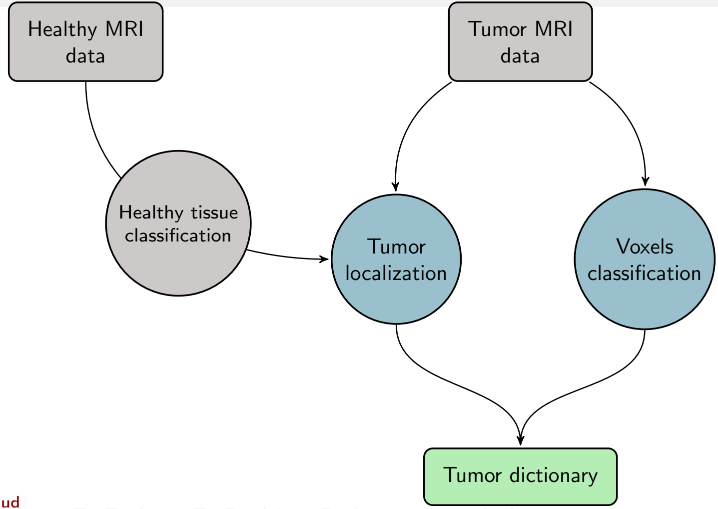
Processing pipeline

Voxels classification



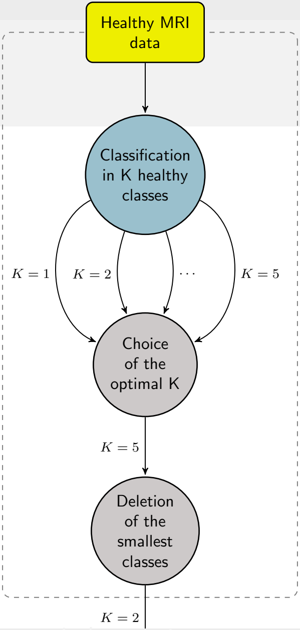
Processing pipeline

Tumor dictionary



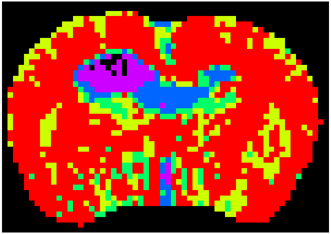
Healthy MRI voxels classification

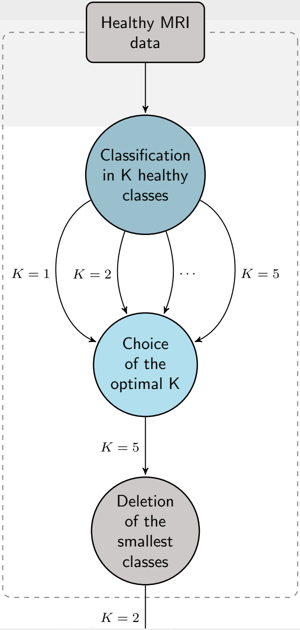
Classification in K healthy classes



Healthy MRI voxels classification

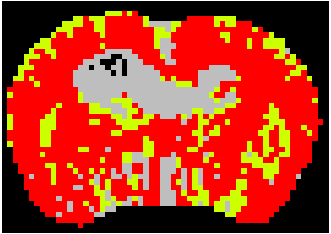
Choice of the optimal K ($K = 5$)

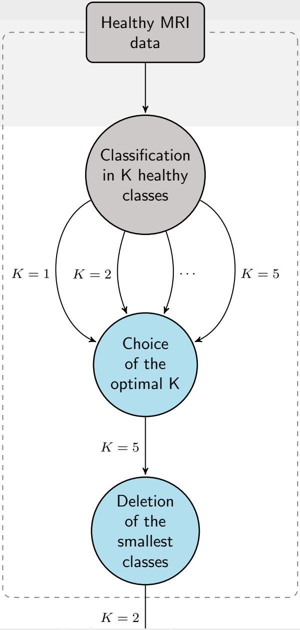




Healthy MRI voxels classification

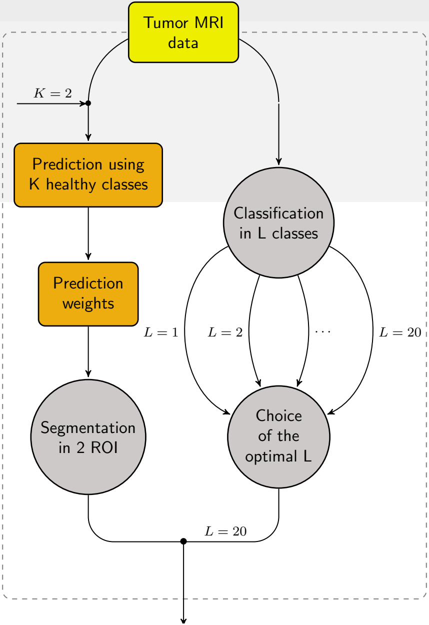
Deletion of the smallest classes ($K = 2$)





Tumor localization & voxels classification

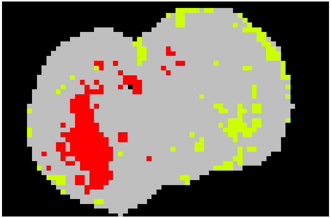
Prediction using K healthy classes to get weights

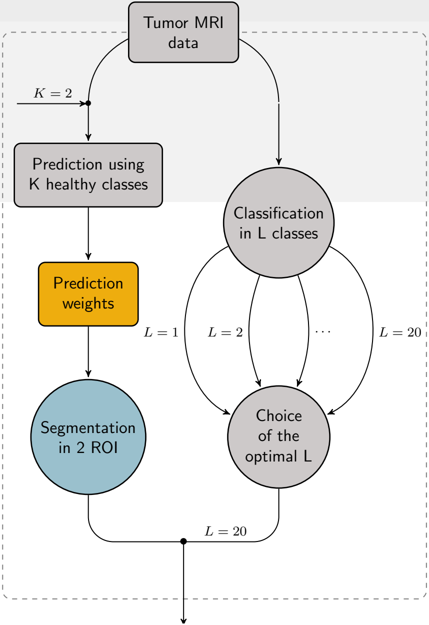


Alexis Arnaud
Supervisors : F. Forbes, E. Barbier, B. Lemasson Laboratories : LJK, INRIA & GIN

Tumor localization & voxels classification

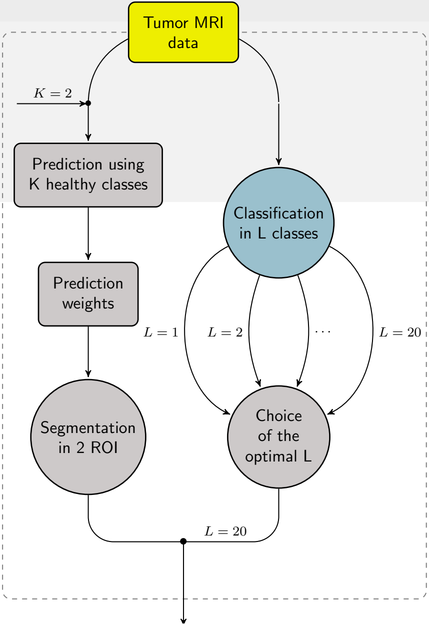
Detection of atypical weight (segmentation in 2 ROI)





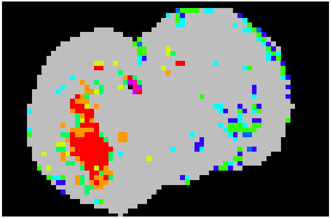
Tumor localization & voxels classification

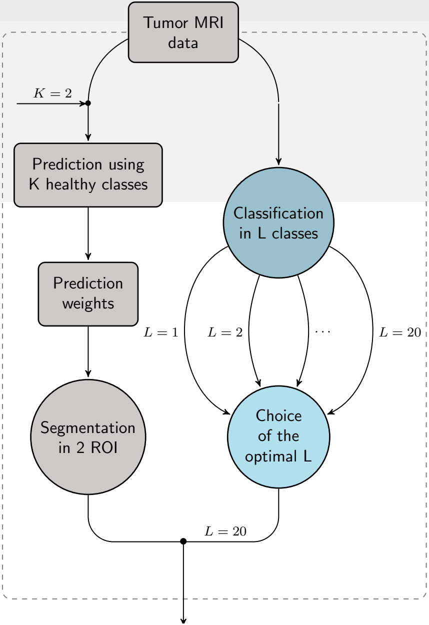
Classification in L tumoral classes



Tumor localization & voxels classification

Choice of the optimal L ($L = 20$)

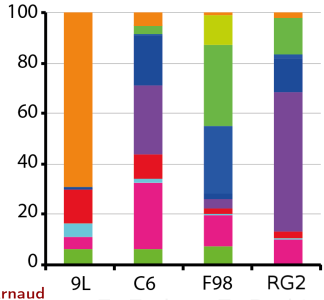


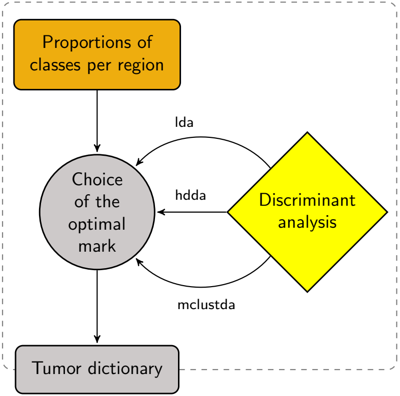


Tumor dictionary

Compute differents features

- Characteristic of the MRI clustering : proportions of classes inside each ROI.
- Discriminant analysis used : lda, hdda, mclustda.

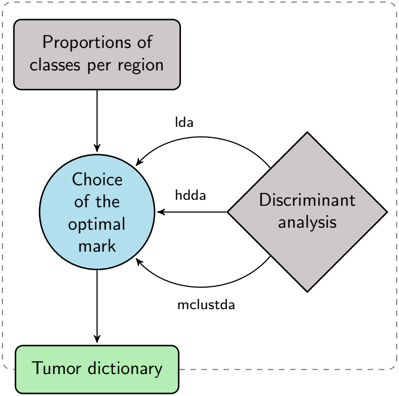




Tumor dictionary

Choice of the optimal mark

- Cross-validation of type : leave-one-out.
- Average classification rate : 70% .



Tumor dictionary

Classification rate

	Prediction							
Truth	**9L**		**C6**		**F98**		**RG2**	
9L outliers (12)	**50**	8.3	.	8.3	.	33	.	.
9L inliers (12)	.	**50**	8.3	.	33	.	8.3	.
C6 outliers (8)	.	25	**62.5**	12.5	.	.	.	.
C6 inliers (8)	.	.	.	**75**	12.5	12.5	.	.
F98 outliers (8)	12.5	12.5	.	.	**75**	.	.	.
F98 inliers (8)	.	.	.	.	.	**100**	.	.
RG2 outliers (7)	.	14.3	.	.	.	.	**85.7**	.
RG2 inliers (7)	14.3	.	.	.	.	.	.	**85.7**

Except for the 9L rats, the classification of tumors is as good as the previous study but with more tumor types and automatic tumor localisation.

Outline

Bayesian extension

Motivations

- more general and flexible model (use of a priori)
- model selection (automatic choice of the class number)

Bayesian extension

Hierarchical view of the current mixture model

$$
\begin{aligned}
\boldsymbol{Y}_n \,|\, \mathrm{Z}_n = k, \boldsymbol{W}_n = \mathbf{w}_n \;;\; \boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k &\sim \mathcal{N}_{\mathrm{M}} \left(\boldsymbol{\mu}_k, \boldsymbol{U}_k \boldsymbol{\Delta}_n^{-1} \boldsymbol{D}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}} \right) \\
\boldsymbol{W}_n = (\boldsymbol{W}_{n1}, \ldots, \boldsymbol{W}_{n\mathrm{M}}) \,|\, \mathrm{Z}_n = k \;;\; \boldsymbol{\nu} &\sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G} \left(\frac{\nu_{km}}{2}, \frac{\nu_{km}}{2} \right) \\
\mathrm{Z}_n \;;\; \boldsymbol{\pi} &\sim \mathcal{M} \left(1, \pi_1, \ldots, \pi_{\mathrm{K}} \right)
\end{aligned}
$$

Bayesian extension

Hierarchical view of the current mixture model

$$\boldsymbol{Y}_n \,|\, \mathrm{Z}_n = k, \boldsymbol{W}_n = \mathbf{w}_n \,;\, \boldsymbol{\mu}_k, \boldsymbol{U}_k, \boldsymbol{D}_k \sim \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{\mu}_k, \boldsymbol{U}_k \boldsymbol{\Delta}_n^{-1} \boldsymbol{D}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}}\right)$$

$$\boldsymbol{W}_n = (\boldsymbol{W}_{n1}, \ldots, \boldsymbol{W}_{n\mathrm{M}}) \,|\, \mathrm{Z}_n = k \,;\, \boldsymbol{\nu} \sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G}\left(\frac{\nu_{km}}{2}, \frac{\nu_{km}}{2}\right)$$

$$\mathrm{Z}_n \,;\, \boldsymbol{\pi} \sim \mathcal{M}(1, \pi_1, \ldots, \pi_{\mathrm{K}})$$

Bayesian extension : a standard prior for $\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k = \boldsymbol{U}_k \boldsymbol{D}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}}$ is :

$$\boldsymbol{\mu}_k \,|\, \boldsymbol{\Sigma}_k \sim \mathcal{N}_{\mathrm{M}}$$

$$\boldsymbol{\Sigma}_k \sim \mathcal{IW}_{\mathrm{M}}$$

Bayesian extension

Hierarchical view of the current mixture model

$$\mathbf{Y}_n \,|\, \mathrm{Z}_n = k, \mathbf{W}_n = \mathbf{w}_n \,;\, \boldsymbol{\mu}_k, \mathbf{U}_k, \mathbf{D}_k \sim \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{\mu}_k, \mathbf{U}_k \boldsymbol{\Delta}_n^{-1} \mathbf{D}_k^{-1} \mathbf{U}_k^{\mathrm{t}}\right)$$

$$\mathbf{W}_n = (\mathbf{W}_{n1}, \ldots, \mathbf{W}_{n\mathrm{M}}) \,|\, \mathrm{Z}_n = k \,;\, \boldsymbol{\nu} \sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G}\left(\frac{\nu_{km}}{2}, \frac{\nu_{km}}{2}\right)$$

$$\mathrm{Z}_n \,;\, \boldsymbol{\pi} \sim \mathcal{M}\left(1, \pi_1, \ldots, \pi_{\mathrm{K}}\right)$$

Bayesian extension : a standard prior for $\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k = \mathbf{U}_k \mathbf{D}_k^{-1} \mathbf{U}_k^{\mathrm{t}}$ is :

$$\boldsymbol{\mu}_k \,|\, \boldsymbol{\Sigma}_k \sim \mathcal{N}_{\mathrm{M}}$$

$$\boldsymbol{\Sigma}_k \sim \mathcal{IW}_{\mathrm{M}}$$

▷ How to get priors on $\mathbf{U}_k, \mathbf{D}_k$?

Bayesian extension

Hierarchical Bayesian mixture model

$$\begin{aligned}
\boldsymbol{Y}_n \,|\, \mathrm{Z}_n = k, \boldsymbol{W}_n = \boldsymbol{w}_n, \boldsymbol{M}_k = \boldsymbol{\mu}_k, \boldsymbol{D}_k = \boldsymbol{d}_k \,;\, \boldsymbol{U}_k &\sim \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{\mu}_k, \boldsymbol{U}_k \boldsymbol{\Delta}_n^{-1} \boldsymbol{d}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}}\right) \\
\boldsymbol{W}_n = (\boldsymbol{W}_{n1}, \ldots, \boldsymbol{W}_{n\mathrm{M}}) \,|\, \mathrm{Z}_n = k \,;\, \boldsymbol{a}, \boldsymbol{b} &\sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G}\left(a_{km}, b_{km}\right) \\
\mathrm{Z}_n \,;\, \boldsymbol{\pi} &\sim \mathcal{M}\left(1, \pi_1, \ldots, \pi_{\mathrm{K}}\right) \\
\boldsymbol{M}_k \,|\, \boldsymbol{D}_k = \boldsymbol{d}_k \,;\, \boldsymbol{U}_k &\sim \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{m}_k, \eta_k^{-1} \boldsymbol{U}_k \boldsymbol{d}_k^{-1} \boldsymbol{U}_k^{\mathrm{t}}\right) \\
\boldsymbol{D}_k = (\boldsymbol{D}_{k1}, \ldots, \boldsymbol{D}_{k\mathrm{M}}) &\sim \bigotimes_{m=1}^{\mathrm{M}} \mathcal{G}\left(\boldsymbol{\alpha}_{km}, \boldsymbol{\beta}_{km}\right)
\end{aligned}$$

Outline

Work in progress

- Validation of the protocol.
- Taking into account spatial dependences using a hidden Markov field.
- Parameters sensitivity analysis.
- Automatic selection of the number of classes.
- Link between histology and automatic tissue characterization.

Bibliography

- **Coquery N, Francois O, Lemasson B, Debacker C, Farion R, Rémy C, et Barbier E (2014)**, *Microvascular MRI and unsupervised clustering yields histology-resembling images in two rat models of glioma*, Journal of Cerebral Blood Flow & Metabolism, volume 34, number 8, 1354–62.

- **Forbes F, et Wraith D (2014)**, *A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweights: Application to robust clustering*, Statistics and Computing, volume 24, number 6, 971–984.

The end

Thank you !

Image credit : dianephotos.deviantart.com