

Mélanges de lois de Student multivariées généralisées : application à la caractérisation de tumeurs par IRM multiparamétrique

Alexis Arnaud^{1,2}, Florence Forbes^{1,2}, Nicolas Coquery^{3,4}, Emmanuel L. Barbier^{3,4}, et Benjamin Lemasson^{3,4}
¹INRIA, ²LJK Université Grenoble Alpes, ³U836 INSERM, ⁴GIN Université Grenoble Alpes ; Grenoble, France

But de l'étude : Dans cette étude nous développons une méthode statistique pour la classification de données IRM multiparamétriques permettant de réaliser un contrôle qualité des données (détection de données atypiques), et de construire un dictionnaire de signatures de tumeurs. Une précédente étude [1] se basait sur le modèle de mélange de lois gaussiennes dans lequel on cherche à rassembler les observations en groupes (classes) issus de lois gaussiennes. Ce modèle est connu pour sa sensibilité aux valeurs atypiques qui peuvent sensiblement dégrader la pertinence des groupes obtenus (eg. Fig.1(b)). Nous proposons donc d'utiliser des lois de Student généralisées [2] qui permettent d'attribuer un poids à chaque observation dans chaque dimension, pour pondérer l'influence des valeurs atypiques sur la forme des groupes. Il en résulte une plus grande flexibilité dans l'ajustement des groupes (Fig.1(a)).

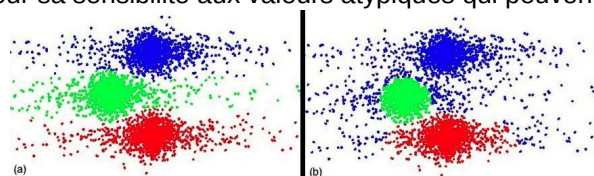


Fig. 1 : classification bivariable en 3 groupes avec lois (a) de Student généralisées [2] et (b) gaussiennes.

Matériel et méthodes: Nous appliquons ce modèle à des IRM de cerveaux de rats présentant quatre modèles de gliomes : 9L (n=5), C6 de deux laboratoires (C6a, n = 6 ; C6b, n = 13), et F98 (n=13). Les images IRM ont été acquises suivant un protocole similaire à [1]. Les paramètres d'intérêt pour chaque rat sont : le coefficient de diffusion apparent (ADC), le volume sanguin cérébral (CBV), le débit sanguin cérébral (CBF), la perméabilité des vaisseaux (AUC), et le taux d'oxygénation des tissus (StO2). Des régions d'intérêt (ROI : 1 tumeur à gauche, 2 saines à droite) ont été manuellement délimitées sur les images pondérée T2, puis reportées sur les cartes paramétriques. Les voxels de toutes les ROI sont rassemblés et classifiés ensemble de façon à regrouper les pixels présentant des caractéristiques micro-vasculaires similaires.

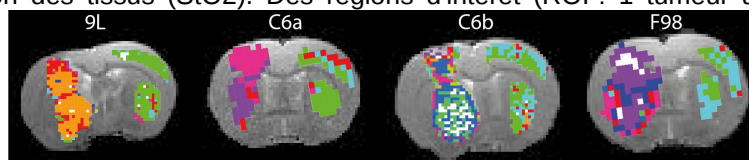


Fig. 2 : classification de cerveaux de rats (10 classes) pour 4 modèles de gliome (9L, C6a, C6b, F98), illustrée sur 1 tranche d'IRM.

Résultats et discussion: La classification des voxels nécessite de fixer le nombre K de classes à estimer. Pour cela nous minimisons le critère d'information bayésien (BIC), ce dernier présentant un minimum marqué contrairement au cas du mélange gaussien [1]. Nous obtenons K=10, ce qui donne une classification (Fig. 2) cohérente avec celle de [1] mais avec des groupes dont les centres sont moins déviés par les valeurs extrêmes et les variances réduites. L'estimation des poids de chaque observation permet d'identifier des clusters où toutes les observations sont proches du centre, ainsi que des clusters avec un nuage d'observations plus éloigné. La classification ainsi réalisée permet également de construire une signature pour chaque ROI en considérant la proportion de chaque cluster au sein d'une ROI (Fig. 3). Les ROI saines présentent une signature similaire au sein de chaque type de rats (« Striat » ; Fig. 3). Cette similarité des ROI saines liée au type permet de détecter des rats atypiques, ce qui est le cas pour l'un des rats (« Outlier striat » ; Fig. 3). Au contraire, les ROI de tumeurs présentent une grande différence entre elles, y compris entre les deux souches C6 (« C6a et C6b » ; Fig. 3). Le pouvoir discriminant des signatures établies est évalué par une procédure leave-one-out : la reconnaissance est quasi parfaite (> 97%) avec une seule tumeur F98 mal-classée en tumeur C6b.

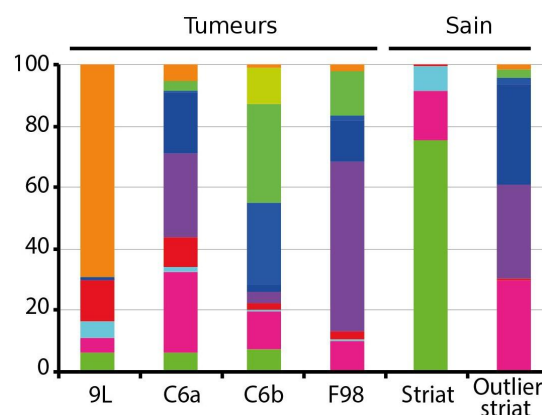


Fig. 3 : signature des ROI représentée comme la moyenne des proportions respectives de chaque cluster dans chaque ROI. La dernière colonne montre la signature de la partie saine d'un rat atypique.

Conclusion: L'utilisation d'un mélange de lois de Student multivariées généralisées donne accès à des informations supplémentaires concernant l'éloignement d'une observation au centre du groupe et ceci pour chaque paramètre. Cela nous a notamment permis de détecter et d'écarter un rat atypique, ainsi que de construire des signatures pour chaque ROI avec un fort pouvoir discriminant pour les modèles de tumeurs, y compris entre les deux gliomes C6 utilisés, ce qui nous laisse penser que ce modèle statistique pourrait détecter des dérives de lignées cellulaires entre laboratoires.

Références:

[1] Coquery N, Francois O, Lemasson B, Debacker C, Farion R, Rémy C, Barbier E. *Microvascular MRI and unsupervised clustering yields histology-resembling images in two rat models of glioma.* J Cereb Blood Flow Metab. 2014 Aug ; 34(8) :1354-62.

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