HAL
open science

Tumor classification and prediction using robust multivariate clustering of multiparametric MRI

Alexis Arnaud, Florence Forbes, Benjamin Lemasson, Emmanuel Luc Barbier

To cite this version:

Alexis Arnaud, Florence Forbes, Benjamin Lemasson, Emmanuel Luc Barbier. Tumor classification and prediction using robust multivariate clustering of multiparametric MRI. International Society for Magnetic Resonance in Medicine, May 2015, Toronto, Canada. hal-01253584

HAL Id: hal-01253584

https://hal.science/hal-01253584v1

Submitted on 11 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

2287

Tumor classification and prediction using robust multivariate clustering of multiparametric MRI

Alexis Arnaud[1,2], Florence Forbes[1,2], Benjamin Lemasson[3,4], and Emmanuel L Barbier[3,4]

[1]*INRIA, Grenoble, -, France,* [2]*LJK, University Grenoble Alpes, Grenoble, -, France,* [3]*U836, INSERM, Grenoble, -, France,* [4]*GIN, University Grenoble Alpes, Grenoble, -, France*

Introduction (Target & Purpose): In neuro-oncology, the use of multiparametric MRI may better characterize brain tumor heterogeneity. To fully exploit multiparametric MRI (e.g. tumor classification), appropriate analysis methods are yet to be developed. In this work, we show on small animals data that advanced statistical learning approaches can help 1) in organizing existing data by detecting and excluding outliers and 2) in building a dictionary of tumor fingerprints from a clustering analysis of their microvascular features.

Methods: Multiparametric MRI were acquired on 4 different brain tumor models: Wistar rats with 9L glioma (n=5) C6 glioma from a first lab (C6a, n=13) ; C6 glioma from a second lab (C6b, n=6) and Fischer rats with F98 glioma (n=13). MRI acquisition was performed at 4.7 T between 21 and 24 days after tumor implantation in the left side of the brain. Acquired MRI maps were: diffusion (ADC), blood volume (CBV), blood flow (CBF), tissue oxygen saturation (StO_2) and vessel permeability (Perm). Three regions of interest (ROI: 2 healthy on the right, 1 tumor on the left) were manually delineated for each rat on T_2w-images, reported on each map and turned into a set of parameter vectors corresponding to the selected voxels. All parameter vectors from all rats were then partitioned into a number of classes K (clusters) with similar MRI characteristics using an Expectation-Maximization (EM) algorithm. K was automatically determined using the Bayesian Information Criterium (BIC). To better accomodate for outlier vectors, we used a new family of multivariate distributions[2] instead of the standard Gaussian distributions used in previous work[1]. This new family is more flexible and has the ability to capture a larger variety of cluster shapes especially in a multivariate setting (Fig. 1). In each ROI of each rat, a signature was then built using the relative proportions of each cluster. The signatures predictive power was assessed with a leave-one-out procedure.

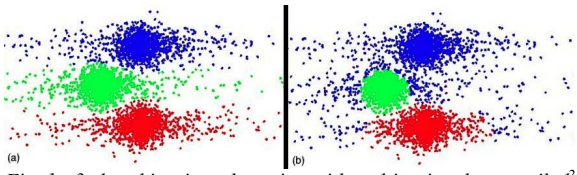

*Fig. 1 : 3 class bivariate clustering with multivariate heavy-tailed[2] (**a**) vs. Gaussian (**b**) distributions. Gaussian mixtures are not able*

Results & Discussion: Our combined EM-BIC procedure led to an optimal number of 10 classes for the 9,030 voxels resulting from merging all ROIs. After assigning a cluster (a color) to each voxel within each ROI, different sub-regions may be identified (Fig. 2). The cluster compositions of healthy ROIs were comparable (mostly green, blue, red; Fig. 2). One animal which presented some atypical cluster composition in these regions was then easily detected and discarded. In contrast, each tumor ROI shows different cluster composition, including the two C6 glioma. To obtain quantitative estimates, respective cluster proportions inside each tumor ROI were computed. For each tumor, its cluster proportions correspond to the tumor signature. Average tumor signatures are illustrated in Fig. 3. A leave-one-out prediction analysis based on the signatures collected in the 36 animals confirmed the visual impression with an almost perfect prediction of the four tumor types: only one F98 type was misclassified as a C6b. The leave-one-out prediction was less efficient for healthy ROIs (between 20 and 92% of prediction). Interestingly, the two C6 cell lines (originating from different labs) showed strong differences. Future work should further evaluate this variability to investigate whether it is inherent to tumors or the simple reflect of population inhomogeneity.

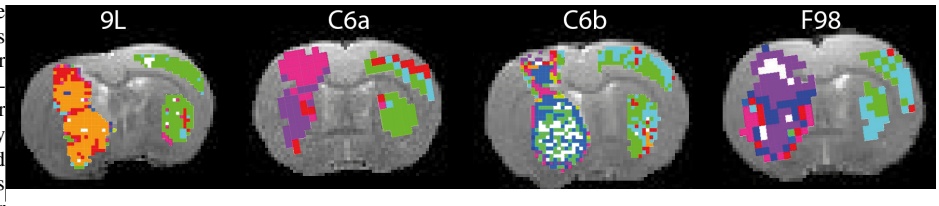

Fig. 2 : Rat brain clustering (10 classes) for 4 glioma models (9L, C6a, C6b, F98) illustrated on 1 slice. Not all 10 classes appear on this slice.

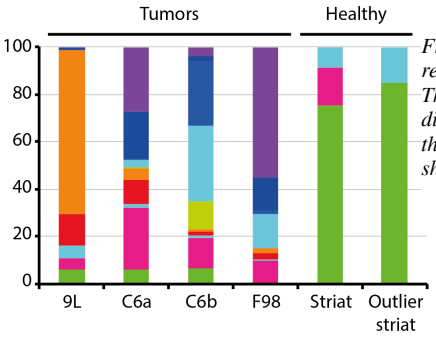

Fig. 3 : Average ROI signatures represented as the relative mean proportions of each cluster in each ROI. The healthy ROI signature (Striat) may easily be distinghuished from the 4 tumor signatures which are themselves very different from each other. The last column shows the healthy ROI signature of an outlier rat.

Conclusion: Advanced statistical clustering approaches are promising tools to better exploit the wealth of MRI information especially on large cohorts and multi-center studies. They offer improved data quality control by allowing automatic outlier detection and improved analysis by identifying discriminative tumor signatures with measurable predictive power. Future work should include the integration in a joint statistical model of both automatic ROI delineation and clustering for whole brain data analysis, with a better use of anatomical information.

References:

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.