



HAL
open science

Evaluation statistique de la segmentation manuelle de données IRM de gliomes diffus de bas grade

Meriem Ben Abdallah, Marie Blonski, Sophie Wantz-Mézières, Yann Gaudeau, Luc Taillandier, Jean-Marie Moureaux

► To cite this version:

Meriem Ben Abdallah, Marie Blonski, Sophie Wantz-Mézières, Yann Gaudeau, Luc Taillandier, et al.. Evaluation statistique de la segmentation manuelle de données IRM de gliomes diffus de bas grade. 18e Colloque Compression et Representation des Signaux Audiovisuels, CORESA 2016, May 2016, Nancy, France. hal-01316788

HAL Id: hal-01316788

<https://hal.science/hal-01316788v1>

Submitted on 17 May 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.

Evaluation statistique de la segmentation manuelle de données IRM de gliomes diffus de bas grade

M. Ben Abdallah¹ M. Blonski^{1,2} S. Mézières³ Y. Gaudeau^{1,4} L. Taillandier^{1,2} J.M. Moureaux¹

¹ Université de Lorraine, Centre de Recherche en Automatique de Nancy (CRAN), CNRS UMR 7039, Faculté de Médecine - Bât D - BP 184, Vandoeuvre-lès-Nancy, 54505, France

² Service de Neuro-Oncologie, Centre Hospitalier Régional Universitaire de Nancy, Avenue du Maréchal Lattre de Tassigny, 54035 Nancy, France

³ Université de Lorraine, Institut de Mathématiques Elie Cartan, INRIA BIGS CNRS UMR 7502, BP 239, F-54506 Vandoeuvre-lès-Nancy Cedex, France

⁴ Université de Strasbourg, 30 Rue du Maire André Trabant, Haguenau, 67500, France

Résumé

Les gliomes diffus de bas grade sont des tumeurs cérébrales primitives rares des adultes. La segmentation manuelle est essentielle pour le suivi des patients atteints de cette tumeur et pour le choix du traitement optimal. Cette méthode étant chronophage, il semble difficile de l'inclure dans la routine clinique. La segmentation automatique apparaît donc comme une solution potentielle pour répondre à cette problématique. Cependant, les algorithmes actuels de segmentation automatique n'ont pas encore prouvé leur efficacité pour les gliomes diffus de bas grade [1][2][3][4][5][6] en raison de la spécificité de ce type de tumeurs. De ce fait, la segmentation manuelle demeure, aujourd'hui, la seule vérité terrain dans ce domaine. Une alternative pour contourner la perte en temps liée à la segmentation manuelle serait de partager la tâche entre différents praticiens, à condition que cette dernière soit reproductible. Le but de notre travail est d'évaluer la reproductibilité de la segmentation manuelle des examens IRM de gliomes diffus de bas grade, en fonction des praticiens, de leur expérience et de leur spécialité. Dans ce travail, nous avons conduit une étude statistique sur les volumes tumoraux d'un panel de 14 experts ayant manuellement segmenté 12 examens IRM de gliomes diffus de bas grade en utilisant le logiciel OsiriX. La plupart des études de segmentation de tumeurs cérébrales publiées mélangent différents types de tumeurs et comparent la segmentation automatique à la segmentation manuelle. Notre étude, au contraire, se focalise uniquement sur les gliomes diffus de bas grade et sur leur segmentation manuelle, car ce sont les plus difficiles à délimiter en raison de leur nature invasive. Une analyse statistique a fourni des résultats prometteurs en démontrant que les facteurs praticien, spécialité médicale et nombre d'années d'expérience n'ont pas d'impact

significatif sur les valeurs moyennes de la variable volume tumoral.

Mots clefs

Segmentation, IRM, Gliomes Diffus de Bas Grade.

1 Introduction

Les Gliomes Diffus de Bas Grade (GDBG) sont des tumeurs cérébrales primitives rares des adultes. Ces tumeurs progressent continuellement au cours du temps pour se transformer en tumeurs de grade supérieur dont la malignité est associée à différentes anomalies neurologiques et dont l'issue est fatale. La taille de la tumeur est l'un des plus importants facteurs pronostiques statiques [7]. La régression linéaire à base de modèles mixtes a mis en avant un taux d'augmentation moyen du diamètre tumoral de 4,1 mm par an pour ce type de tumeurs [8]. La stratégie thérapeutique est basée sur une approche personnalisée en plusieurs étapes et avec, à long terme, une adaptation au cours du temps en fonction des changements volumétriques et cliniques. La chirurgie fonctionnelle précoce est généralement le premier traitement proposé lorsque cela est possible. La chimiothérapie peut être utilisée en tant que traitement adjuvant mais aussi, parfois, en traitement néoadjuvant avant la résection chirurgicale. La radiothérapie est habituellement réservée en cas de progression après une chimiothérapie de tumeurs non résécables ou au moment de la transformation anaplasique [9]. Pour le suivi des patients, il est essentiel d'appréhender l'évolution volumétrique dans des conditions cliniques habituelles (pendant les consultations) afin d'adapter de façon optimale le traitement en temps réel [10]. La comparaison qualitative simple de deux examens IRM séparés de 4 à 6 mois d'intervalle ne permet pas habituellement d'objectiver la crois-

sance. Il a été proposé à l'origine de mesurer les 3 plus grands diamètres dans les 3 plans de l'espace, D1, D2, D3, et ensuite d'extrapoler le volume avec la formule suivante : $D1 \cdot D2 \cdot D3 / 2$ [11]. Une méthode de segmentation manuelle sur logiciel a été développée dans [12] et est devenue, depuis, la technique de référence pour la majorité des experts. Comme cette technique est chronophage, une segmentation par différents cliniciens permettrait d'accélérer la prise en charge thérapeutique des patients. Cependant, à notre connaissance, la reproductibilité de la segmentation manuelle sur les images IRM de GDBG n'a pas encore été évaluée [13]. En effet, les principales recherches actuelles englobent plusieurs types de tumeurs cérébrales dans la même étude et se concentrent plutôt sur une comparaison de performance entre segmentation automatique et manuelle. Si la segmentation automatique peut être d'un grand intérêt [14] [15], nous affirmons que, dans le cas des GDBG, la segmentation manuelle reste non seulement la seule vérité terrain mais également la meilleure méthode pour déterminer le volume de ces tumeurs pour la majorité des équipes spécialisées. En effet, la segmentation automatique ne semble pas encore être assez fiable pour distinguer les anomalies de signal de la tumeur d'autres causes d'anomalies de signal (modifications post-chirurgicales ou post-radiothérapie, leucoencéphalopathie de diverses étiologies, etc.). Comme la segmentation manuelle est chronophage, la méthode moins précise des 3 diamètres est généralement préférée pour le suivi du volume tumoral en pratique hospitalière régulière. Le travail que nous proposons ici aborde la question de la reproductibilité de la segmentation manuelle en étudiant l'incidence du praticien sur l'estimation du volume du GDBG. En effet, ce dernier peut fortement influencer le choix d'une thérapie ainsi que les instants de son début et de son arrêt. Ces répercussions motivent la conduite d'une étude subjective de la cohérence de la segmentation manuelle au sein d'un groupe d'experts de GDBG. Une telle cohérence est la clé de la fiabilité et de la reproductibilité du diagnostic clinique et, par conséquent, de la sélection de la thérapie appropriée. Le reste du papier est organisé comme suit. Dans le paragraphe 2, la méthodologie et les outils utilisés dans cette étude sont décrits. Dans le paragraphe 3, les techniques statistiques d'évaluation sont détaillées. Le paragraphe 4 présente les résultats de l'étude statistique. Le paragraphe 5 résume ces résultats et discute de leur conséquences pour la pratique médicale.

2 Matériels et méthodes

Les tests subjectifs permettent d'évaluer la qualité des images et des vidéos, comme par exemple en compression de données à différents taux de compression [16][17]. Ces tests prennent en considération les pratiques médicales les plus courantes et sont réalisés dans un environnement strictement contrôlé. Les résultats des tests subjectifs sont quantifiés par des métriques objectives et se basent, pour leur interprétation, sur une vérité terrain pré-définie. Dans cette étude, la moyenne des volumes a été choisie comme

vérité terrain en raison de l'absence d'une vérité terrain absolue. Un expert neuroradiologue a sélectionné 12 examens IRM de 9 patients diagnostiqués avec un GDBG et n'ayant subi aucun traitement préalable. Cet expert ne fait, bien entendu, pas partie du panel de testeurs pour la suite. Les données ont toutes été fournies en coupes axiales, avec une pondération en FLAIR, sauf un examen, qui présentait une pondération en T2. Les examens en pondération FLAIR sont actuellement utilisés en pratique clinique pour le suivi des GDBG et en l'absence de cette séquence IRM, on recourt aux séquences en T2. L'IRM en pondération FLAIR, comme dans l'exemple de la Figure 1, fait apparaître la tumeur en hypersignal, sans le Liquide Cérébro-Spinal (LCS), et ce, contrairement à la séquence en T2, qui rehausse le LCS en plus de la tumeur, augmentant ainsi la difficulté de la segmentation. Par ailleurs, dans les examens sélectionnés, il y avait 3 examens IRM en Cube FLAIR et 9 examens en acquisition FLAIR usuelle. L'étude de la reproductibilité de la segmentation manuelle a été réalisée dans le Living Lab PROMETEE (PeRceptiOn utilisateur pour les usages du MultimEdia dans les applications mEdicalEs : User perception for multimedia medical usages).¹ Cette plate-forme, située à TELECOM Nancy, école d'ingénieurs de l'Université de Lorraine, est une plate-forme d'innovation permettant l'étude et la bonne gestion de la qualité technique de vidéos et d'images pour un usage médical. Cette plate-forme est également dédiée au post-traitement d'images et vidéos pour une utilisation multimédia dans le domaine médical. Elle est bien équipée et aménagée, offrant un environnement conforme aux conditions générales de visualisation pour ce genre de tests, telles que fixées par la recommandation ITU-BT.500-13 [18]. L'éclairage de la pièce a été contrôlé afin de ne pas produire de réflexion sur l'écran. Un panel de 14 experts a segmenté manuellement l'ensemble des données avec OsiriX comme illustré dans l'exemple de la Figure 1.

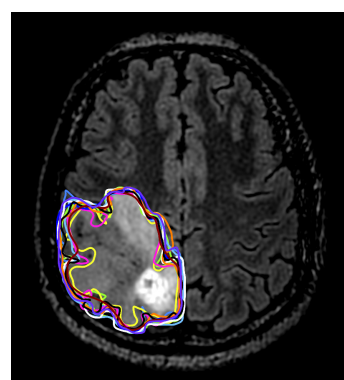


Figure 1 – Exemple de la segmentation manuelle d'une coupe IRM avec OsiriX. Chaque courbe colorée correspond à la segmentation effectuée par un participant.

OsiriX est un logiciel libre de visualisation de fichiers

1. <http://telecomnancy.univ-lorraine.fr/fr/recherche/living-lab>.

Tableau 1 – *Distribution des participants par spécialité médicale.*

Spécialité médicale	Neurologie	Radiologie	Radiothérapie
Nombre de participants	6	4	3

Tableau 2 – *Distribution des participants par années d'expérience.*

Années d'expérience]0 ; 10]]10 ; +∞[
Nombre de participants	8	5

Dicom sous Apple Macintosh [19] [20]. C'est l'un des meilleurs logiciels de segmentation manuelle d'accès libre, qui est largement utilisé au sein de la communauté de neuro-oncologie. La version 32 bits d'OsiriX a été adoptée pour cette étude. Les participants ont commencé par un test visuel sur une tablette afin de détecter les participants ayant des problèmes d'acuité visuelle. Ensuite, ils ont segmenté un examen d'apprentissage, qui n'a pas été inclus dans les résultats de l'étude, de manière à se familiariser avec l'outil de segmentation OsiriX. L'instruction consistait à délimiter manuellement les contours de la tumeur sur des coupes contenant une prise de contraste liée à un GDBG. Pour être en accord avec la pratique médicale, le fenêtrage radiologique et le nombre de coupes à segmenter n'ont pas été précisés. Les participants ont, d'abord, segmenté la moitié de l'ensemble des examens, ont pris 5 minutes de pause, puis ont complété la segmentation de la deuxième moitié des données. A la fin, ils ont rempli un questionnaire sur leur spécialité médicale et leurs années d'expérience depuis le résidanat. Suite aux premiers tests de cohérence, il s'est avéré que l'un des participants avait des résultats incohérents. Ainsi, tous les résultats décrits par la suite sont basés sur les évaluations réalisées sur les 13 participants cohérents. Pour l'étude de la variabilité introduite par la spécialité médicale sur la tumeur, trois catégories ont été définies : neurologues, radiologues et radiothérapeutes. Pour les années d'expérience, deux groupes ont été créés :]0 ; 10] and]10 ; +∞[. La distribution des spécialités médicales et des années d'expérience est répertoriée, respectivement, dans les tableaux 1 et 2. A l'issue du test, les contourages manuels ont été sauvegardés et le volume tumoral, pour chaque examen IRM, a été calculé sous OsiriX en utilisant la méthode de reconstruction des triangulations de Delaunay.

3 Analyse statistique

Les données retenues pour l'étude comportent 12 volumes tumoraux pour chacun des 13 participants : $(x_{i,j})_{i=1..13,j=1..12}$. Le but de l'analyse statistique est d'étudier la variabilité introduite par le facteur praticien sur la variable volume tumoral afin d'examiner l'influence du praticien sur les volumes tumoraux obtenus. Cette étude

visée, également, à analyser la relation entre la spécialité médicale du participant ainsi que ses années d'expérience et les volumes tumoraux. Pour l'étude de la variabilité introduite par le praticien, une analyse de la variance à un facteur (ANOVA) [21] a été appliquée aux volumes tumoraux. Afin de quantifier statistiquement la variabilité introduite par les années d'expérience et la spécialité médicale sur les volumes tumoraux, un volume standard, $y_{i,j}$, a été calculé comme suit :

$$y_{i,j} = \left(\frac{x_{i,j} - \bar{x}_j}{\sigma_j} \right) \quad (1)$$

où \bar{x}_j est le volume moyen, et σ_j est l'écart-type d'un volume. En centrant $x_{i,j}$ autour de la valeur moyenne des volumes pour un ensemble de données et en divisant par son écart-type, on tient compte de la difficulté de segmentation. L'écart-type, σ_{y_i} , de $y_{i,j}$, a ensuite été calculé et un test exact de Fisher [22] a été appliqué sur σ_{y_i} pour les deux études.

Afin d'évaluer la variabilité inter-observateurs, le coefficient de variation (COV) [2][4] en volume a été utilisé. Ce coefficient mesure la variation en volume des objets segmentés, et est défini comme suit :

$$COV_j = \frac{\sigma_j}{\bar{x}_j}$$

Une autre métrique qui est utilisée pour évaluer la variabilité inter-participants est l'indice d'accord (AI) [3]. Cette métrique donne l'accord inter-participants, par paires de participants, pour chaque volume $j = 1, \dots, 12$:

$$AI_{(i,i'),j} = 1 - \frac{2|x_{i,j} - x_{i',j}|}{x_{i,j} + x_{i',j}}$$

pour toutes les paires de participants $(i, i'); i \neq i'; i, i' \in \{1, \dots, 13\}$.

4 Résultats

L'analyse statistique a été réalisée avec le logiciel R. Tout d'abord, les volumes des différents participants ont été tracés avec la moyenne des volumes qui est la vérité terrain sélectionnée. Sur la Figure 2, nous observons de faibles variations de volume, même pour les examens IRM en Cube FLAIR, qui sont censés être plus difficiles à segmenter que des examens FLAIR usuels, et l'ensemble des courbes se confond bien avec la courbe de la vérité terrain. Ce premier résultat est confirmé par la boîte à moustaches sur la Figure 3, où la dispersion autour de la moyenne et de la médiane des volumes tumoraux par examen est faible. Cela suggère que les praticiens ont tendance à segmenter les GDBG de manière similaire. Afin de confirmer ce résultat visuel, une ANOVA a été réalisée sur l'ensemble des données. Avec un seuil de signification de 5%, nous avons pu conclure que le facteur praticien n'a pas d'influence significative sur les valeurs moyennes de la variable volume

tumoral.

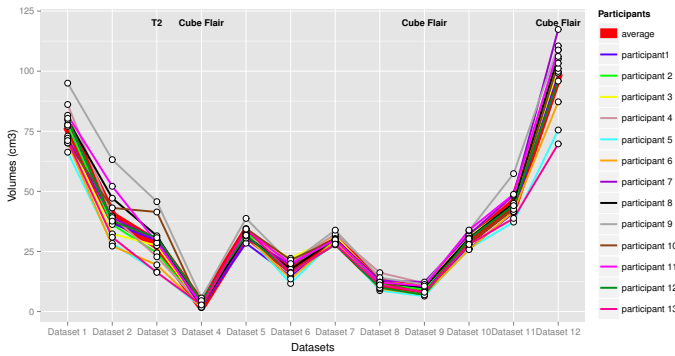


Figure 2 – Variation du volume tumoral pour tous les participants et de la moyenne des volumes en fonction des examens IRM.

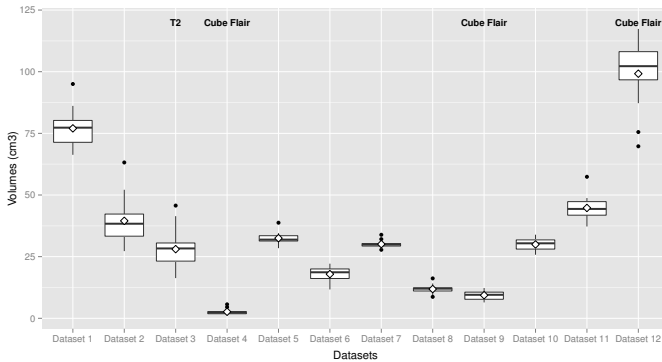


Figure 3 – Boîtes à moustache des volumes tumoraux pour tous les examens.

En ce qui concerne la variabilité introduite par la spécialité médicale sur la variable volume tumoral, nous avons appliqué le test exact de Fisher avec un seuil de signification de 5% et nous avons conclu, pour une valeur-p égale à 0,604, que la spécialité médicale n'avait pas d'impact significatif sur l'estimation de la valeur du volume tumoral. Le tableau 3 confirme cette assertion pour la spécialité médicale. Cela montre que la moyenne des AI et que la moyenne des COV sont très proches entre les différentes spécialités. En outre, l'application des tests de Kolmogorov-Smirnov sur le COV par paires de groupes (avec un seuil de signification de 5%) a confirmé ce résultat. A noter que les volumes tumoraux varient de $1,67 \text{ cm}^3$ à $117,35 \text{ cm}^3$ pour les différents examens. On a, donc, une grande variabilité de la taille du volume. Le COV est évidemment plus sensible aux faibles volumes.

Quant à la variabilité produite par les années d'expérience sur la variable volume tumoral, l'application du test exact de Fisher donne une valeur-p de 0,8961, ce qui indique clairement que l'expérience n'a pas d'influence significative sur le volume segmenté. Ce résultat est, de plus, confirmé

Tableau 3 – COV et AI par spécialité médicale.

Spécialité médicale	Neurologie	Radiologie	Radiothérapie
COV (Moyenne±ET)	17.99 ±12.44	16.56 ±10.11	14.48 ±12.32
AI (Moyenne±ET)	0.74 ±0.28	0.73 ±0.27	0.74 ±0.27

Tableau 4 – COV et AI par années d'expérience.

Années d'expérience]0 ; 10]]10 ; +∞[
COV (Moyenne±ET)	16.58 ±11.09	14.86 ±11.88
AI (Moyenne±ET)	0.75 ±0.28	0.73 ±0.27

par les résultats du tableau 4.

5 Discussions et conclusions

Dans ce travail, nous avons évalué la reproductibilité de la segmentation manuelle du volume tumoral sur des IRM pour différents praticiens, en fonction de leur expérience et de leur domaine d'expertise. Cette étude montre que, en moyenne, ni le praticien médical, ni la spécialité ou l'expérience n'a d'impact significatif sur le volume tumoral. Ce dernier résultat est plutôt surprenant : en effet, on s'attendrait plutôt à ce que l'expérience soit un facteur discriminant. Une autre observation surprenante est que les plus grandes différences en volume ont été remarquées sur les 3 premiers examens, qui sont censés être faciles à contourner. La répartition autour de la médiane de leurs boîtes à moustache (voir Figure 3) est grande par rapport, par exemple, aux examens 4 et 11, qui sont des examens Cube Flair et qui sont censés être plus compliqués à segmenter. Cela pourrait s'expliquer par la nouveauté de l'outil de segmentation utilisé, OsiriX, pour certains participants. Donc, sur les premiers examens, les participants n'avaient pas encore maîtrisé cet outil. Parmi les examens en Cube Flair, l'examen 14 semble avoir été plus difficile à contourner que les examens 4 et 11. Ceci pourrait s'expliquer par l'effet de la fatigue à la fin du test. Mais cela ne semble pas affecter le résultat global de l'étude. Sur la base de plusieurs critères couramment utilisés dans la littérature, tels que le coefficient de variation et l'indice d'accord, l'analyse statistique réalisée pour cette étude a démontré l'absence d'impact, à la fois de la spécialité médicale et des années d'expérience, sur le volume tumoral segmenté, et ce, quelle que soit la difficulté du jeu de données (Cube vs IRM classique). Ce résultat encourageant favorise la collaboration interdisciplinaire entre les cliniciens, en particulier pour les consultations alternées qui sont très fréquentes entre neurochirurgiens et neuro-oncologues. Et même si ce résultat doit être confirmé par des études plus larges, il ouvre la porte à des perspectives intéressantes dans le contexte difficile de la prise en charge des GDBG, où la segmentation automatique ne semble pas encore être en mesure d'offrir une solution totalement fiable. Comme conséquence directe des résultats de nos travaux, le processus de segmen-

tation manuelle pourrait être accéléré, car de nombreux cliniciens seraient en mesure de contourner les GDBG de différents patients, même ceux qu'ils ne suivent pas. Le suivi de l'évolution de la tumeur serait amélioré puisque moins de temps serait consacré par les praticiens à la segmentation manuelle et plus de temps pourrait être dédié aux décisions cliniques concernant le traitement approprié à prescrire à différents stades de la maladie.

Remerciements

Nous adressons nos remerciements à tous les cliniciens du CHRU de Nancy et du CHR de Metz qui ont participé à cette étude. A savoir, nous tenons à remercier : Pr Serge Bracard, Pr Luc Taillandier, Dr Marie Blonski, Dr Basile Wittwer, Dr Guillaume Vogin, Dr Christian Delgoffe, Dr Claire Griffaton-Taillandier, Dr Marie-Alexia Ottenin, Dr Sophie Planel, Dr Fabien Rech, Dr Valérie Bernier, Camille Dahan, Dr Emanuelle Schmitt, Dr Lavinia Jager Simon et Dr Philippe Quetin.

Références

- [1] G. P. Mazzara, R. P. Velthuizen, J. L. Pearlman, H. M. Greenberg, et H. Wagner. Brain tumor target volume determination for radiation treatment planning through automated mri segmentation. *Int. J. Radiation Oncology Biol. Phys.*, 59, No. 1 :300–312, May 2004.
- [2] C. Weltens, J. Menten, M. Feron, E. Bellon, P. Demaerel, F. Maes, W. Van den Bogaert, et E. van der Schueren. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiotherapy & Oncology*, 60 :49–59, Juillet 2001.
- [3] K. Xie, J. Yang, Z. G. Zhang, et Y. M. Zhu. Semi-automated brain tumor and edema segmentation using mri. *European Journal of Radiology*, 56 :12–19, Oct 2005.
- [4] M. R. Kaus, S. K. Warfield, A. Nabavi, P. M. Black, F. A. Jolesz, et R. Kikinis. Automated segmentation of mr images of brain tumors. *Radiology*, 218 :586–591, Feb 2001.
- [5] K. E. Emblem, B. Nedregaard, J. K. Hald, T. Nome, P. Due-Tonnessen, et A. Bjornerud. Automatic glioma characterization from dynamic susceptibility contrast imaging : brain tumor segmentation using knowledge-based fuzzy clustering. *Journal of Magnetic Resonance Imaging*, 30 :1–10, Jul 2009.
- [6] I. Njeh, L. Sallemi, I. ben Ayed, K. Chtourou, S. Lehericy, D. Galanaud, et A. Ben Hamida. 3d multimodal mri brain glioma tumor and edema segmentation : a graph cut distribution matching approach. *Computerized Medical Imaging and Graphics*, 40 :108–119, March 2015.
- [7] L. Capelle, D. Fontaine, E. Mandonnet, L. Taillandier, J. L. Golmard, L. Bauchet, J. Pallud, P. Peruzzi, M. H. Baron, M. Kujas, J. Guyotat, R. Guillemin, M. Fréna, S. Taillibert, P. Colin, V. Rigau, F. Vandenbos, C. Pinelli, H. Duffau, et French Réseau d'Étude des Gliomes.. Spontaneous and therapeutic prognostic factors in adult hemispheric world health organization grade ii gliomas : a series of 1097 cases : clinical article. *Journal of Neurosurgery*, 118 :1157–1168, June 2013.
- [8] E. Mandonnet, J.-Y. Delattre, M. L. Tanguy, K. R. Swanson, A. F. Carpentier, H. Duffau, P. Cornu, R. Van Effenterre, E. C. Jr Alvord, et L. Capelle. Continuous growth of mean tumor diameter in a subset of grade ii gliomas. *Annals of Neurology*, 53 :524–528, April 2003.
- [9] H. Duffau et L. Taillandier. New concepts in the management of diffuse low-grade glioma : Proposal of a multistage and individualized therapeutic approach. *Neuro-Oncology*, August 2014.
- [10] J. Pallud, L. Taillandier, L. Capelle, D. Fontaine, M. Peyre, F. Ducray, H. Duffau, et E. Mandonnet. Quantitative morphological magnetic resonance imaging follow-up of low-grade glioma : a plea for systematic measurement of growth rates. *Neurosurgery*, 71 :729–739, sep 2012.
- [11] J. Pallud, M. Blonski, E. Mandonnet, E. Audureau, D. Fontaine, N. Sanai, L. Bauchet, P. Peruzzi, M. Fréna, P. Colin, R. Guillemin, V. Bernier, M. H. Baron, J. Guyotat, H. Duffau, L. Taillandier, et L. Capelle. Velocity of tumor spontaneous expansion predicts long-term outcomes for diffuse low-grade gliomas. *Neuro-Oncology*, 15 :595–606, Feb 2013.
- [12] E. Mandonnet, J. Pallud, D. Fontaine, L. Taillandier, L. Bauchet, P. Peruzzi, J. Guyotat, V. Bernier, M. H. Baron, H. Duffau, et L. Capelle. Inter- and inpatient comparison of who grade ii glioma kinetics before and after surgical resection. *Neurosurgical Review*, 33 :91–96, 2010.
- [13] M. C. Chamberlain. Is the volume of low-grade glioma measurable and is it clinically relevant? *Neuro-Oncology*, 16 :1027–1028, June 2014.
- [14] Z. Akkus, J. Sedlar, L. Coufalova, P. Korfiatis, T. L. Kline, J. D. Warner, J. Agrawal, et B. J. Erickson. Semi-automated segmentation of pre-operative low grade gliomas in magnetic resonance imaging. *Cancer Imaging*, 15(1) :1–10, 2015.
- [15] N. Porz, S. Bauer, A. Pica, P. Schucht, J. Beck, R. K. Verma, et R. Wiest. Multi-modal glioblastoma segmentation : Man versus machine. *PLoS ONE*, 9(5), May 2014.
- [16] Y. Gaudeau, J. Lambert, N. Labonne, et J.-M. Moureaux. Compressed image quality assessment : application to an interactive upper limb radiology atlas. Dans *IEEE International Conference on Image Processing, ICIP 2014*, Paris, France, Octobre 2014.

- [17] A. Chaabouni, Y. Gaudeau, J. Lambert, J.-M. Mouraux, et P. Gallet. Subjective and objective quality assessment for h264 compressed medical video sequences. Dans *4th International Conference on Image Processing Theory, Tools and Applications, IPTA'14*, Paris, France, Octobre 2014.
- [18] ITU-R Rec-BT.500. Recommendation 500-13, methodology for the subjective assessment of the quality of television pictures. 2012.
- [19] A. Rosset, L. Spadola, et O. Ratib. Osirix : An open-source software for navigating in multidimensional dicom images. *Journal of Digital Imaging*, 17 :205–216, Septembre 2004.
- [20] Osirix downloads. <http://www.osirix-viewer.com/Downloads.html>.
- [21] G. E. P. Box, W. G. Hunter, et J. S. Hunter. *Statistics for Experimenters : an introduction to Design, Data Analysis, and Model Building*. Wiley Series in Probability and Mathematical Statistics, 1978.
- [22] D. G. Altman. *Practical statistics for medical research*. Chapman & Hall, 1991.