

Segmentation de tumeur cérébrale à partir d'IRM multi-séquences par apprentissage de noyaux multiples

Naouel BOUGHATTAS^{1,2}, Maxime BERAR², Kamel HAMROUNI¹, Su RUAN²

¹Laboratoire Signal Image et Technologies de l'Information (LR-SITI)
1 Université de Tunis El Manar, 1068, Tunis, Tunisie

²Laboratoire d'Informatique, du Traitement de l'Information et des Systèmes (LITIS)
1 Université de Rouen, 22, Boulevard GAMBETTA, 76183 ROUEN, France
naouel.boughattas@gmail.com, maxime.berar@univ-rouen.fr
kamel.hamrouni@enit.rnu.tn, su.ruan@univ-rouen.fr

Résumé – Plusieurs types d'image IRM (Imagerie par Résonance Magnétique) peuvent être utilisées par les médecins pour faire un diagnostic sur les tumeurs cérébrales et estimer leurs volumes. La fusion de ces différentes sources fournit une aide pour la segmentation de la tumeur et l'estimation de son volume. Nous proposons une méthode de segmentation basée sur la classification permettant de prendre en compte les différentes caractéristiques de chaque type d'image. Cette méthode est basée sur un algorithme SVM multi-noyaux qui sélectionne les caractéristiques les plus pertinentes en associant un noyau pour chaque caractéristique extraite. Le poids du noyau reflète la pertinence de la caractéristique correspondante. La sélection des caractéristiques et le réglage du paramètre de régularisation du SVM se font d'une manière automatique. La méthode a été appliquée sur 7 patients. Nous effectuons dans cet article une comparaison avec une méthode récente de la littérature qui a été elle-même comparée avec différentes méthodes pour montrer sa compétitivité [2]. Nos expérimentations montrent une bonne performance de notre méthode.

Abstract – To make a diagnosis on brain tumours, several types of MRI image (Magnetic Resonance Imaging) can be used. The fusion of these various sources helps medical experts to segment the tumor and estimate its volume. We propose a segmentation method based on classification using extracted features from each type of image. This method is based on a multiple kernel learning SVM algorithm which selects the most relevant features by associating a kernel for each extracted feature. The weight of the kernel reflects the relevance of the corresponding feature. The feature selection and the selection of the regularization parameter of the SVM are made automatically. The method has been tested on 7 patients. We show in this article a comparison with another recent method which was compared with various methods to show its effectiveness [2]. Experiments show a good performance of our method.

1 Introduction

Les tumeurs cérébrales doivent être détectées dès leurs premiers stades afin de pouvoir contrôler leur évolution durant tout le traitement thérapeutique et offrir une chance de guérison aux patients. Les experts du domaine prennent beaucoup de temps pour segmenter manuellement la tumeur et quantifier l'évolution de son volume, avec à la clé une variabilité inter-experts. Un système de segmentation capable de fournir des résultats quantitatifs sur la taille de la tumeur dans les images IRM est nécessaire.

Ce travail renvoie le problème de segmentation de la tumeur à un problème de classification [2,3,4,7,9], où chaque pixel est associé soit à la classe des tissus sains soit à la classe des tissus pathologiques. La classification se fait généralement à l'aide d'un Séparateur à Vaste Marge (SVM) permettant de construire un modèle de prédiction qui assigne, pour un nouveau point l'une des deux classes

selon son emplacement par rapport à un hyperplan séparateur déterminé durant la phase d'apprentissage effectuée à partir d'un petit nombre de points.

Dans [2], les auteurs présentent une méthode efficace pour segmenter la tumeur à partir de la fusion de 3 séquences IRM (T2, Rho : Densité de Proton, et FLAIR) en s'appuyant sur un classifieur SVM. La fusion des modalités s'opère en extrayant sur chaque séquence un ensemble de caractéristiques calculé à travers une fenêtre centrée sur chaque pixel de chacune de ces 3 séquences et en les regroupant dans un même vecteur. Les caractéristiques les plus significatives sont ensuite sélectionnées à l'aide d'un critère de séparabilité de classes dans un espace noyau. Cette méthode donne de bons résultats mais nécessite une étape lourde de validation pour la fixation manuelle d'un seuil lors la sélection des caractéristiques puis lors de l'entraînement du classifieur SVM. Autant d'étapes de validation à réaliser à partir d'une base d'apprentissage limitée par le temps

disponible de l'expert à étiqueter des pixels représentatifs. Pour remédier à ces problèmes, nous proposons d'utiliser un algorithme d'apprentissage des noyaux multiples pour classifieur SVM (Multiple Kernel Learning-SVM) [5,1] qui réalise simultanément la sélection de poids associés à des noyaux et l'entraînement du classifieur. Chacune des caractéristiques précédentes pouvant être associée à un ou plusieurs noyaux, le problème de sélection de caractéristiques peut être transformé en problème de sélection de noyaux. Plusieurs formulations et implémentations de MKL-SVM ont été proposées dans la littérature. Nous avons choisi d'utiliser l'implémentation SimpleMKL [5] pour sa simplicité et son efficacité.

Notre contribution est la mise en œuvre d'une méthode globale qui partant de quelques points étiquetés par l'expert réalise la segmentation du volume d'images de manière complètement automatique basée sur cet algorithme et des noyaux associés à des caractéristiques extraites sur chaque séquence.

2 Les noyaux multiples

Si les données sont linéairement non séparables dans leur espace de description, l'idée est de projeter ces données dans un espace de redescription de très grande dimension (éventuellement de dimension infinie). Dans cet espace de redescription, presque tous les problèmes non linéairement séparables deviennent linéairement séparables. On peut alors y apprendre facilement un classifieur de type SVM.

2.1 Le classifieur SVM

Etant donné un ensemble de points d'apprentissage $\{\mathbf{x}_i\}_{i=1}^l$, et leurs étiquettes de décision (tumeur ou non tumeur) $\{y_i\}_{i=1}^l$, la fonction de décision du SVM est donnée par la formule (1) où k est un noyau défini positif donné (par exemple le noyau Gaussien $k(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}}$) et où les α_i^* sont des paramètres appris lors de la phase d'apprentissage :

$$f(\mathbf{x}) = \sum_{i=1}^l \alpha_i^* k(\mathbf{x}, \mathbf{x}_i), \quad (1)$$

Le problème d'optimisation d'apprentissage sous sa forme duale est donné à travers la formule (2) avec $C \geq 0$ le paramètre de régularisation relatif aux points mal classés.

$$\max_{\alpha_i} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \quad (2)$$

s.t. $0 \leq \alpha_i \leq C, \forall i.$

2.2 L'apprentissage par noyaux multiples

Si les données à traiter sont issues de sources hétérogènes, l'idée est de sélectionner une ou plusieurs fonctions noyaux

k_m de $\mathcal{X}_i \times \mathcal{X}_i$ dans \mathbb{R} pour chaque type de caractéristiques (intensité, mesures statistiques, coefficients d'ondelettes, ...) dans chaque modalité pour chaque pixel i étiqueté par y_i dans $\{-1, 1\}$ et d'exploiter ainsi leur diversité et leur complémentarité. À chacun de ces noyaux est associé un poids positif d_m , leur combinaison linéaire donne alors un noyau, utilisable par un classifieur SVM.

L'apprentissage par noyaux multiples (MKL) a été introduit par Lanckriet et al. [1] et a été développé pour déterminer les poids associés à chaque noyau. La fonction de décision est alors de la forme

$$f(\mathbf{x}) = \sum_{i=1}^l \sum_{m=1}^p \alpha_i^* d_m^* k_m(\mathbf{x}, \mathbf{x}_i). \quad (3)$$

La forme duale du problème d'optimisation MKL [5] introduit un paramètre λ lié aux poids des noyaux.

$$\begin{aligned} \max_{\alpha_i, \lambda} \quad & \sum_i \alpha_i - \lambda \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C, \quad \forall i \\ & \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j k_m(\mathbf{x}_i, \mathbf{x}_j) \leq \lambda, \quad \forall m. \end{aligned}$$

La formulation primale du problème SVM-MKL proposée par [5] reprend le problème SVM traditionnel où la fonction noyau est alors un ensemble de fonctions f_m associées à chaque noyaux k_m .

$$\begin{aligned} \min_{\{f\}_m, \xi, d} \quad & \frac{1}{2} \sum_m d_m \|f_m\|_{\mathcal{H}_n} + C \sum_i \xi_i \quad (4) \\ \text{s.t.} \quad & y_i \sum_m f_m(x_i) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i, \\ & \sum_m d_m = 1, \quad d_m \geq 0, \quad \forall m \end{aligned}$$

où $\|\cdot\|_{\mathcal{H}_m}$ est la norme de la fonction noyau dans son espace de Hilbert, C est le paramètre de régularisation du problème SVM et les $\{\xi_i\}$ la violation de la marge. Pour forcer une sélection de noyaux parcimonieuse, une contrainte est imposée sur le poids des noyaux ($\sum_m d_m = 1, d_m \geq 0, \forall m$).

Dans ce qui suit, nous utilisons l'algorithme SimpleMKL [5] basé sur la résolution alternée du problème primal.

3 Segmentation par MKL

La méthode proposée se déroule en 5 étapes :

- étiquetage de quelques points de chaque classe par l'expert,
- extraction des caractéristiques associées à chaque pixel de l'apprentissage,
- réglage automatique du paramètre de régularisation C par résolutions successives du problème SVM-MKL,

- extraction des caractéristiques retenues sur l'ensemble des pixels,
- et enfin application du classifieur et du noyau combiné à l'ensemble des caractéristiques.

Le schéma représenté par la figure 1 résume l'ensemble de ces étapes.

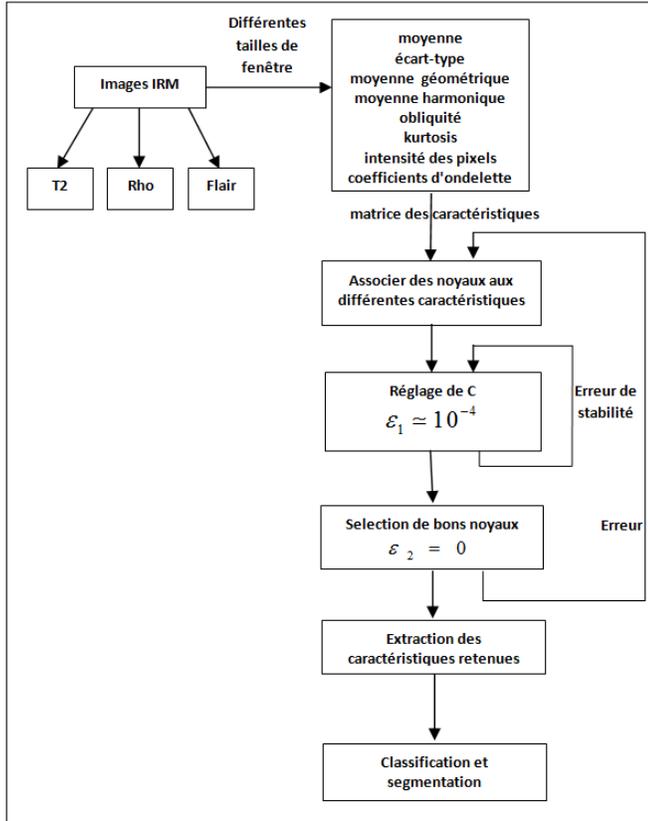


FIG. 1: Schéma de la méthode proposée

3.1 Extraction de caractéristiques

Les caractéristiques les plus répandues pour segmenter la tumeur dans les images IRM sont : l'intensité des pixels, les mesures statistiques [8] et les coefficients d'ondelettes [2] calculés sur une fenêtre dont la taille optimale est inconnue.

Nous proposons de construire notre base en utilisant l'intensité des pixels, la moyenne, l'écart-type, la moyenne géométrique, la moyenne harmonique, l'obliquité, le kurtosis et les coefficients d'ondelette calculés pour chaque séquences IRM (T2, Rho et FLAIR) sur différentes tailles de fenêtres (5, 7, 9, 11 et 13). Il existe évidemment des redondances dans ces caractéristiques et la taille totale du vecteur de caractéristiques en chaque point est grande (≈ 3000).

A chaque caractéristique, on associe 3 noyaux Gaussiens avec 3 paramètres σ différents (1, 10 et 100 fois l'écart-

type). Le nombre total des noyaux est important et vaut 360. L'algorithme MKL devra donc sélectionner les bons noyaux parmi ces 360 noyaux.

3.2 Détermination du paramètre de régularisation du SVM

Le paramètre de régularisation C (Eq. 4) est ajusté automatiquement sans avoir recours à une étape de validation croisée difficile pour si peu d'exemples d'apprentissage: en commençant par une très petite valeur, nous augmentons régulièrement C jusqu'à atteindre des poids des noyaux stables (seuil $\varepsilon_1 = 10^{-4}$, condition $\sum_m (d_m^{t+1} - d_m^t)^2 < \varepsilon_1$). Avec la valeur obtenue de C , nous examinons le vecteur de poids des noyaux générés et nous éliminons les noyaux insignifiants en utilisant un deuxième seuil ε_2 ($d_m \geq \varepsilon_2, \forall m$).

Pour les noyaux restants, nous pouvons raffiner l'ensemble des valeurs possibles de σ (i.e. $\sigma/2$, σ et 2σ). Cette phase peut être itérée jusqu'à la stabilité des noyaux sélectionnés et permet de régler automatiquement les paramètres des noyaux Gaussiens.

4 Évaluation et résultats

Notre algorithme a été testé sur 7 patients, pour lesquels on dispose de 3 volumes IRM de 24 coupes de taille 512x512. Les images IRM des patients ont été acquises à l'aide d'une machine 1.5T GE (General Electric Co.) utilisant un écho de spin rapide axial (FSE :Fast Spin Echo) en pondération T2, Rho et FLAIR. Toutes les images ont été recalées à l'aide de l'outil SPM [10], qui peut être téléchargé à partir du site (<http://www.fil.ion.ucl.ac.uk/spm/>).

Dans la phase d'apprentissage, 60 points (30 appartenant à la classe tumeur et 30 appartenant à la classe non tumeur) ont été choisis manuellement sur une seule coupe.

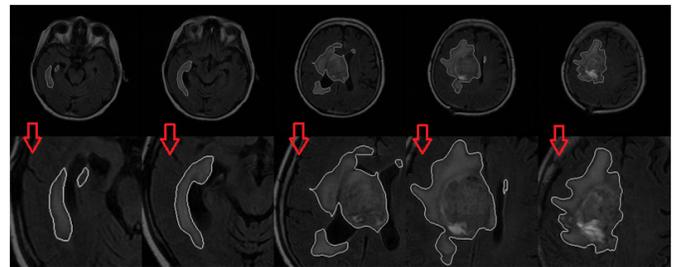


FIG. 2: Résultats de segmentation sur le patient 3 (coupes 7, 8, 13, 15 et 17.)

Les noyaux sélectionnés correspondent aux caractéristiques suivantes : les intensités des pixels (35%) et les coefficients d'ondelette (65%) connus par leur potentiel dans la description des images [2]. Les mesures statistiques sont généralement redondantes et ne sont pas choisies avec le

TAB. 1: Évaluation et comparaison de notre méthode et de [2] (valeur moyenne sur l'ensemble des patients).

	SVM+	MKL
TP(%)	89,15 ± 9,92	95,47 ± 1,75
FP (%)	0,40 ± 0,21	0,33 ± 0,15
FN (%)	10,85 ± 9,92	4,53 ± 1,75
Erreur Totale (%)	11.22 ± 9,72	4.86 ± 1,70

critère de parcimonie. Les modalités T2 et Flair plus contrastées que les images PD sont conservées à part égales. Le volume segmenté grandit avec l'augmentation du paramètre de régularisation. La valeur optimale de C est celle qui trouve un compromis entre les erreurs de classification et la marge dure du classifieur SVM. La figure 2 illustre les résultats de segmentation obtenus avec notre méthode sur 5 coupes du patient 3 appartenant à un seul examen IRM.

L'évaluation de nos résultats en termes d'erreurs de classification est présentée dans le tableau 1 en comparant avec les tracés manuels des experts. Ce tableau montre que les erreurs de classification données par notre algorithme sont plus faibles que celles obtenues par Zhang [2] dont nous avons repris les caractéristiques. Notre méthode semble surtout générer des faux négatifs, ce qui entraîne une légère sous-estimation de la taille globale de la tumeur. Cependant, cela permet de ne pas traiter de tissus sains mal étiquetés et est donc préférable à un système générant beaucoup de faux positifs.

5 Conclusion

Une nouvelle approche de segmentation pour la tumeur cérébrale à partir des images IRM multi-séquences, basée sur un algorithme de noyaux multiples, a été présentée dans ce papier.

Notre méthode attribut un noyau spécifique pour chaque caractéristique, ce qui permet d'exploiter la diversité et la complémentarité de toutes les données utilisées et de sélectionner automatiquement les caractéristiques pertinentes.

Etant donné que le nombre des points fournis par l'expert est réduit, nous proposons d'étendre notre méthode pour des problèmes d'apprentissage semi-supervisé, pour améliorer le processus d'apprentissage.

References

[1] G. Lanckriet, N. Cristianini, L. El Ghaoui, P. Bartlett, et M. Jordan. *Learning the kernel matrix with semi-definite programming*. Journal of Machine Learning Research, vol. 5, pp. 27-72, 2004.

[2] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, et Y. Zhu. *Kernel feature selection to fuse multispectral mri images for brain tumour segmentation*. Computer

Vision and Image Understanding, vol. 115, pp. 256-269, 2011.

- [3] D. Zikic, B. Glocker, E. Konukoglu, J. Shotton, A. Criminisi, D. H. Ye, C. Demiralp, O. M. Thomas, T. Das, R. Jena, et S. J. Price. *in Context-sensitive Classification Forests for Segmentation of Brain Tumor Tissues*. MICCAI, 2012.
- [4] H. Khotanlou, O. Colliot, J. Atif, et I. Bloch. *3d brain tumour segmentation in mri using fuzzy classification, symmetry analysis and spatially constrained deformable models*. Fuzzy Sets and Systems, vol. 160, no. 10, pp. 1457-1473, 2009.
- [5] A. Rakotomamonjy, F. Bach, S. Canu, et Y. Grandvalet. *Simplemkl*. Journal of Machine Learning Research, vol. 9, pp. 2491-2521, 2008.
- [6] L. Wang. *Feature selection with kernel class separability*. IEEE Transactions on pattern analysis and machine intelligence, vol. 30, no. 9, pp. 1534-1546, 2008.
- [7] N. Sharma, A. K. Ray, S. Sharma, K. Shukla, S. Pradhan, et L. M. Aggarwal. *Segmentation and classification of medical images using texture-primitive features : Application of bam-type artificial neural network*. Journal of Medical Physics, pp. 119-126, 2008.
- [8] A. Mekhmoukh, K. Mokrani, et N. Nasri. *A modified fuzzy c means algorithm for noisy images segmentation : Application to mri images*. International Journal of Research and Reviews in Soft and Intelligent Computing, vol. 1, no. 4, 2011.
- [9] Aydin Ulas, Mehmet Gonen, Umberto Castellani, Vittorio Murino, Marcella Bellani, Michele Tansella, et Paolo Brambilla. *A Localized MKL Method for Brain Classification with Known Intra-class Variability*. MICCA, LNCS 7588, pp. 152-159, 2012.
- [10] J. Ashburner et al. *SPM Manual*. June, 2008.