

Segmentation semi-automatique de tumeurs du foie en TDM dynamique pour l'estimation du taux de nécrose

Pierre-Henri CONZE¹, François ROUSSEAU², Vincent NOBLET¹, Fabrice HEITZ¹, Riccardo MEMEO³, Patrick PESSAUX³

¹ICube UMR 7357, Université de Strasbourg, CNRS, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg, France
300 boulevard Sébastien Brant, CS 10413, F-67412 Illkirch Cedex, France

²Institut Mines-Télécom, Télécom Bretagne, INSERM, LATIM UMR 1101
Technopôle Brest-Iroise, CS 83818, 29238 Brest Cedex 3, France

³Pôle Hépato-Digestif, Nouvel Hôpital Civil, Institut Hospitalo-Universitaire de Strasbourg
1 place de l'Hopital, 67000 Strasbourg, France

conze@unistra.fr

Résumé – Les traitements pré-opératoires du carcinome hépato-cellulaire (CHC) retardent la progression des tumeurs par nécrose. Il a été démontré que le taux de nécrose était à ce jour le meilleur critère pour l'évaluation de la réponse à ces traitements. Nous proposons une méthode semi-automatique à faible interaction estimant le taux de nécrose à partir d'images tomographiques (TDM) multi-phases. Celle-ci segmente les tissus de foie sain, tumoraux actifs et nécrosés *via* une forêt d'arbres décisionnels appliquée à des régions sémantiques et exploitant des descripteurs multi-phases basés cluster. L'évaluation confirme l'importance des informations dynamiques extraites à l'échelle des régions.

Abstract – Pre-operative locoregional treatments (PLT) delay the tumor progression by necrosis for patients with hepato-cellular carcinoma (HCC). To accurately assess the PLT response, we address the estimation of liver tumor necrosis (TN) from dynamic contrast-enhanced CT scans. To overcome inter-expert variability induced by visual qualitative assessment, we propose a semi-automatic method that requires weak interaction efforts to segment parenchyma, tumoral active and necrotic tissues. It applies random decision forests on semantic regions and involves robust multi-phase cluster-based features. Quantitative results confirms the benefits of exploiting dynamic cluster-based information.

1 Introduction

Le carcinome hépato-cellulaire (CHC) est la troisième cause la plus fréquente de décès par cancer [1]. L'enjeu est de trouver un traitement adapté à chaque patient en intégrant des données individuelles (age, comorbidités, stade), médicales (nombre, taille, position des tumeurs) et collectives (attente avant transplantation). Avec un taux de survie à 5 ans supérieur à 50% [1], les résections (RH) et transplantations hépatiques (TH) sont des options thérapeutiques viables. Cependant, pour des tumeurs non-résecables ou avant TH, la progression des tumeurs peut être retardée par des traitements pré-opératoires tels que la chimio-embolisation transartérielle, l'ablation radio-fréquence ou le *Sorafénib* [2], un anti-cancéreux oral. Ceux-ci mènent à la destruction des tumeurs par nécrose. Les critères standards d'estimation de la réponse à ces traitements (RECIST, mRECIST, EASL ou WHO) ne permettent cependant pas une évaluation fiable [2]. Des études récentes ont prouvé que le taux de nécrose tumorale est plus fortement corrélé au taux de survie après TH [3] et RH [4]. Celui-ci est actuellement évalué visuellement à partir soit d'images tomographiques (TDM) ou par résonance magnétique (IRM) soit de coupes histologiques.

Pour éviter la forte variabilité inter-expert de l'évaluation quantitative visuelle, nous abordons l'estimation du taux de né-

crose par la biais d'une méthode semi-automatique. Calculer le taux de nécrose nécessite la localisation des tumeurs et la segmentation des zones actives et nécrosées les constituants (Fig. 2c). Puisque le CHC se caractérise par une hypervascularisation intense au temps artériel et un lavage tardif au temps portal en réponse à l'injection de produits de contraste [1], nous proposons d'exploiter des images *multi-phases* issues de la modalité TDM dynamique. Celle-ci résulte, pour chaque examen, en une série d'images acquises avant injection et après injection aux temps artériel, veineux précoce et veineux tardif.

Les images *multi-phases* ont récemment été exploitées en segmentation de tumeurs du foie *via* des descripteurs de dynamique temporelle avec *lignes de niveaux* [5] ou *coupes de graphes* [6]. L'extraction de tumeurs a également fait l'objet d'études récentes d'un point de vue interactif par apprentissage supervisé se basant sur des descripteurs spatiaux [7, 8]. Dans notre contexte, l'interaction utilisateur apparaît nécessaire car le rehaussement artériel et le lavage portal dépendent du protocole d'injection et de la cinétique de l'agent de contraste.

Nous proposons d'appliquer une forêt aléatoire à des régions sémantiques et d'y exploiter des descripteurs multi-phases basés cluster. En basant l'apprentissage et la prédiction sur des clusters, nous discriminons les tissus sur la base d'informations dynamiques et spatiales robustes tout en limitant l'interaction.



FIGURE 1 – Exemple d’images TDM multi-phases recalées (\tilde{I}_{SI} , \tilde{I}_{AR} , I_{VP} et \tilde{I}_{VT}) avec masque de segmentation du foie associé.

2 Méthode

2.1 Approche existante : forêt d’arbres décisionnels *simple-phase* basée *voxel*

En segmentation interactive d’images, l’opérateur définit un ensemble de K voxels labélisés $\mathcal{S} = \{\mathbf{v}_k, c(\mathbf{v}_k)\}_{k \in \{1, \dots, K\}}$ où $c(\mathbf{v}_k) \in \{c_1, \dots, c_N\}$ est l’étiquette assignée par l’opérateur au voxel \mathbf{v}_k parmi N étiquettes. Cet ensemble \mathcal{S} est utilisé par les méthodes d’apprentissage telles que les forêts d’arbres décisionnels [9] pour construire un modèle permettant de prédire l’étiquette $c(\mathbf{v})$ de chaque voxel de test \mathbf{v} . Une forêt d’arbres décisionnels se décompose en T arbres indépendants faits de nœuds internes et terminaux. Chaque nœud interne affecte les voxels à ses fils gauche et droite *via* un test binaire effectué sur un sous-ensemble $\hat{\theta}(\mathbf{v})$ des descripteurs $\theta(\mathbf{v})$ associés à \mathbf{v} :

$$h(\mathbf{v}, \theta) = \begin{cases} \text{vrai, si } \tau_1 < \hat{\theta}(\mathbf{v}) < \tau_2 \\ \text{faux, sinon} \end{cases} \quad (1)$$

où $\hat{\theta}(\mathbf{v})$ est comparé aux seuils τ_1 et τ_2 . Les approches utilisant des forêts d’arbres décisionnels pour la segmentation de tumeurs [7, 8] ne reposent que sur une seule image et exploitent ainsi des descripteurs $\theta(\mathbf{v})$ non dynamiques, dits *simple-phase*.

Lors de l’apprentissage, chaque arbre prend en entrée les données d’apprentissage \mathcal{S} et optimise ses propres nœuds internes ($\{\tau_1, \tau_2, \hat{\theta}(\mathbf{v})\}$) afin d’en déduire les tests binaires les plus discriminants. Après optimisation, chaque nœud terminal l_t de l’arbre t reçoit une partition \mathcal{S}_{l_t} de \mathcal{S} et génère une distribution de probabilités pour l’ensemble des étiquettes, $p_{l_t}(c_i | \mathcal{S}) \forall i \in \{1, \dots, N\}$. Pour prédire l’étiquette $c(\mathbf{v})$ d’un voxel de test \mathbf{v} , celui-ci est propagé dans chacun des T arbres optimisés sur la base des descripteurs *simple-phase* $\theta(\mathbf{v})$. Il en résulte :

$$c(\mathbf{v}) = \arg \max_{c_i} \sum_{t=1}^T p_{l_t}(c(\mathbf{v}) = c_i | \mathcal{S}) \quad (2)$$

2.2 Stratégie proposée : forêt d’arbres décisionnels *multi-phases* basée *cluster*

Extension aux régions. Dans l’approche traditionnelle, la segmentation souffre d’un manque de cohérence spatiale car les voxels sont traités indépendamment. Des descripteurs sont parfois dédiés au contexte spatial [7, 8] mais leur étendue reste limitée. De plus, l’apprentissage basé voxel nécessite une forte interaction pour obtenir un ensemble d’apprentissage assez représentatif. Afin de réduire l’effort d’interaction et d’introduire de la cohérence spatiale, nous proposons d’appliquer la forêt d’

arbres décisionnels non pas à des voxels mais à des clusters d’apprentissage et de test.

Extension au multi-phases. Puisque le diagnostic du CHC repose sur l’association hypervascularisation artérielle et lavage portal [1], nous proposons de tirer profit d’images TDM *multi-phases*. Lors d’un examen, un agent de contraste est injecté et des images sont acquises avant injection (SI) et après injection aux temps artériel (AR), vénéux précoce (VP) et tardif (VT). Chaque examen conduit à la création d’un ensemble d’images $\mathcal{S}_I = \{I_{SI}, I_{AR}, I_{VP}, I_{VT}\}$. Ces images sont ensuite recalées par rapport à I_{VP} (Fig. 1) car I_{VP} présente de plus grands contrastes inter-tissus que les autres acquisitions. Le recalage est non rigide difféomorphe [10] pour compenser les déformations liées à la respiration. Nous faisons l’hypothèse qu’un masque de segmentation du foie (Fig. 1) est disponible au temps VP.

Protocole proposé. Notre méthode consiste à classer les tissus *via* une forêt d’arbres décisionnels¹ *multi-phases* basée *cluster* (FA-MpCl) ce qui permet d’en déduire le taux de nécrose. En pratique, 3 étiquettes sont considérées : tissus sains (parenchyme), tumoraux actifs et nécrosés.

Le volume du foie est tout d’abord sur-segmenté par une décomposition en supervoxels 3D SLIC² [11]. A partir des voxels appartenant au masque du foie dans I_{VP} , nous obtenons ainsi un ensemble de K_R clusters $\mathbf{R} = \{\mathbf{r}_i\}_{i \in \{1, \dots, K_R\}}$ (Fig. 2a). SLIC intègre l’information spatiale en favorisant l’agrégation de voxels voisins dont les intensités dans I_{VP} sont proches.

L’étape suivante a trait à la sélection interactive des clusters d’apprentissage. Au lieu de considérer un grand nombre de voxels, l’opérateur n’a qu’à sélectionner et étiqueter un sous-ensemble de \mathbf{R} (Fig. 2b). Une telle interaction nous semble plus adaptée à la pratique clinique du fait de sa simplicité et de sa possible intégration dans un protocole d’apprentissage multi-examens. Il en résulte un ensemble de clusters d’apprentissage $\mathcal{S} = \{\mathbf{r}_j, c(\mathbf{r}_j)\}_{j \in \{1, \dots, K_R\}}$ associant aux clusters \mathbf{r}_j les étiquettes $c(\mathbf{r}_j) \in \{c_1, c_2, \dots, c_N\}$. Une forêt d’arbres décisionnels est ensuite construite sur la base de \mathcal{S} . L’apprentissage optimise chacun des nœuds internes par rapport aux descripteurs *multi-phases* basés *cluster* $\theta(\mathbf{r}_j)$ calculés pour chaque \mathbf{r}_j .

Les clusters de test \mathbf{r} ($\mathbf{r} \in \mathbf{R} \setminus \mathcal{S}$) sont ensuite propagés dans la forêt d’arbres décisionnels afin d’obtenir une prédiction de leur étiquette $c(\mathbf{r})$ (Fig. 2c) sur la base de leurs propres descripteurs *multi-phases* basés *cluster*. Enfin, le taux de nécrose τ est calculé comme le rapport du volume des zones nécrosées sur le volume total de la tumeur (ie, zones actives et nécrosées).

1. implémentation *scikit-learn* : <http://scikit-learn.org/>
2. implémentation *scikit-image* : <http://scikit-image.org/>

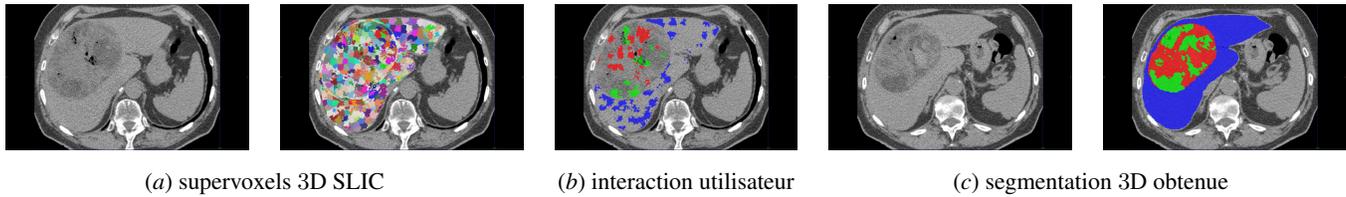


FIGURE 2 – Illustration de la méthode de segmentation semi-automatique *via* une forêt d’arbres décisionnels *multi-phases* basée *cluster* (FA-MpCl). Celle-ci segmente au sein du volume hépatique les tissus sains (bleu), tumoraux actifs (rouge) et nécrosés (vert) (c) en basant l’interaction utilisateur (b), l’apprentissage ainsi que la prédiction sur des régions sémantiques 3D (a).

TABLE 1 – Descripteurs *multi-phases* basés *cluster* exploités au sein de la forêt d’arbres décisionnels.

Catégories	Descripteurs <i>multi-phases</i> basés <i>cluster</i>	Nombre
Intensité spatiale	intensité moyennée + écart-type	4 + 4
Gradient spatial	gradient moyenné + écart-type	4 + 4
Multi-phase	pic de rehaussement (PE)	1
	aire sous la courbe de rehaussement (AUC)	1
	différences inter-phases $\Delta_{VP/AR}$, $\Delta_{VT/VP}$	2

Descripteurs *multi-phases* basés *cluster*. La qualité de la classification dépend de la faculté des descripteurs à discriminer les différents tissus. À chaque région sémantique sont assignés 20 descripteurs *multi-phases* basés *cluster*, regroupés en 3 catégories (Tab. 1). Les deux premières traitent des caractéristiques spatiales en intensité et en gradient à l’échelle des clusters et quantifient l’intensité intrinsèque, l’homogénéité, la présence de texture ainsi que la répartition des échelles de texture. La troisième combine plusieurs critères d’intensité *multi-phases* : les critères dynamiques utilisés dans [5, 6] incluant le pic de rehaussement (PE) et l’aire sous la courbe (AUC) auxquels sont ajoutées les différences inter-phase $\Delta_{VP/AR}$ et $\Delta_{VT/VP}$ (Fig. 3). Ces descripteurs dynamiques discriminent les clusters sur la base de leur rehaussement artériel et lavage portal.

3 Résultats

Protocole d’évaluation. Le protocole est testé sur 7 examens. Pour chacun, 6 coupes axiales régulièrement réparties sont choisies pour couvrir l’étendue spatiale des tumeurs. 4 experts en chirurgie hépato-digestive ont délimité les tissus sains, actifs et nécrosés pour chaque coupe. Ces masques sont fusionnés avec STAPLE [12] menant à un total de 42 coupes véritées-terrain.

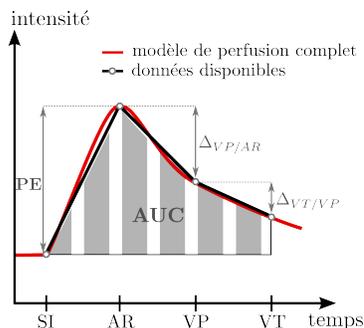


FIGURE 3 – Illustration des descripteurs *multi-phases*.

Chaque examen est traité indépendamment. Pour chacun d’eux, nous identifions parmi l’ensemble des clusters 3D SLIC ceux dont l’intersection avec les coupes véritées-terrain manifeste une étiquette vérité-terrain dominante (consensus sur au moins 95% des voxels). Cette étiquette est assignée au cluster 3D correspondant. 1/3 des clusters identifiés sont choisis aléatoirement (pour simuler l’interaction utilisateur) et sont dédiés à l’apprentissage de la forêt d’arbres décisionnels *multi-phase* basée *cluster* (FA-MpCl). Une fois entraînée, la forêt sert à l’étiquetage des clusters restants. Nous présentons des comparaisons avec les approches *simple-phase* basée *voxel* (FA-SpVx), *simple-phase* basée *cluster* (FA-SpCl) et *multi-phases* basée *voxel* (FA-MpVx). Chaque forêt contient $T=100$ arbres.

Pour rendre la comparaison possible, les voxels d’apprentissage pour *Sp/MpVx* sont ceux constituant les clusters d’apprentissage sélectionnés pour *Sp/MpCl*. Les descripteurs utilisés pour chaque stratégie diffèrent. *MpCl* assigne à chaque cluster les descripteurs de Tab. 1 tandis que *SpCl* n’en utilise qu’une liste réduite : les descripteurs *multi-phases* sont écartés et les intensités et gradients mis en jeu ne se basent que sur I_{VP} . Pour les méthodes basées *voxel*, *MpVx* assigne à chaque voxel \mathbf{v} et pour chaque phase les intensités et gradients intrinsèques et moyennés dans un bloc 3^3 centré autour de \mathbf{v} auxquels sont ajoutés PE, AUC, $\Delta_{VP/AR}$ et $\Delta_{VT/VP}$. *SpVx* n’exploite que les informations d’intensité et de gradient de I_{VP} .

SpVx, *SpCl*, *MpVx* et *MpCl* sont évalués, par rapport aux masques véritées-terrain, en termes d’erreur en taux de nécrose $\Delta\tau = |\tau - \tau_{\text{expert}}|$ et de coefficients *DICE* pour les tissus sains, actifs et nécrosés : $DICE_s$, $DICE_a$ et $DICE_n$. Pour être moins sensible aux aspects aléatoires des forêts décisionnelles, les résultats de chaque examen sont moyennés sur 10 réalisations.

Résultats. Tab. 2 présente pour les 4 méthodes les $\Delta\tau$ et coefficients *DICE* moyennés sur toute la base de données. Des

TABLE 2 – Comparaison des forêts d’arbres décisionnels (FA) *simple-phase* basées *voxel* (*SpVx*), *simple-phase* basées *cluster* (*SpCl*), *multi-phases* basées *voxel* (*MpVx*) ainsi que *multi-phases* basées *cluster* (*MpCl*) en termes d’erreur en taux de nécrose et de *DICE* moyennés sur l’ensemble des données.

scores	FA-SpVx	FA-SpCl	FA-MpVx	FA-MpCl
$\Delta\tau$	6, 40 ± 2, 85	9, 13 ± 4, 78	6, 60 ± 3, 32	5,26 ± 3,90
$DICE_a$	54, 3 ± 17, 2	65, 8 ± 15, 3	65, 5 ± 12, 4	74,4 ± 12,6
$DICE_n$	65, 0 ± 21, 6	63, 8 ± 25, 8	71, 8 ± 17, 6	71,9 ± 19,5
$DICE_s$	80, 5 ± 13, 1	89, 7 ± 4, 90	87, 4 ± 9, 00	93,3 ± 3,08

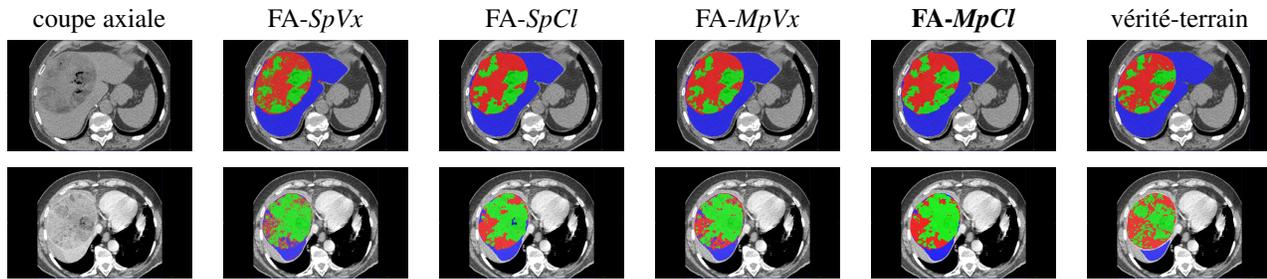


FIGURE 4 – Exemples de segmentation de tumeurs hépatiques par le biais de forêts d’arbres décisionnels (FA) *simple-phase* basées *voxels* (*SpVx*), *simple-phase* basées *clusters* (*SpCl*), *multi-phases* basées *voxels* (*MpVx*) ainsi que *multi-phases* basées *clusters* (*MpCl*) avec les masques vérités-terrain associés. Les tissus de foie sain, actifs et nécrosés sont en bleu, rouge et vert.

résultats visuels sont fournis Fig. 4 avec vérité-terrain associée. L’étude comparative révèle que la meilleure méthode est FA-*MpCl* avec des *DICEs* de 74,4, 71,9 et 93,3 pour les tissus sains, actifs et nécrosés ainsi que le $\Delta\tau$ le plus faible (5,26). FA-*MpVx* obtient un $DICE_n$ équivalent mais de moins bons $DICE_a$, $DICE_s$ et $\Delta\tau$. De plus, Fig. 4 illustre une plus forte cohérence spatiale avec FA-*MpCl*. Enfin, FA-*MpCl* nécessite une interaction moins importante que pour FA-*MpVx*.

La comparaison des approches simple/multi-phases confirme que la prise en compte de données multi-phases améliore les résultats avec des gains en *DICEs* de l’ordre de 8,6 (11,2), 8,1 (6,8) et 3,6 (6,9) resp. en basé cluster et voxel. Les différences significatives entre FA-*SpCl* et FA-*MpCl* justifient l’utilisation de descripteurs multi-phases robustes quand les forêts aléatoires sont appliquées à des clusters : une erreur sur un seul cluster peut avoir un fort impact. Enfin, soulignons que la variabilité inter-experts moyenne calculés à l’aide des vérités-terrain est de $21,2 \pm 15,2$ sur le taux de nécrose ! Cela justifie la nécessité d’une approche semi-automatique dédiée à l’estimation du taux de nécrose pour réduire le biais induit par l’évaluation visuelle.

4 Conclusion

Nous présentons une méthode semi-automatique de segmentation de tumeurs pour l’évaluation du taux de nécrose dans le cadre de traitements pré-opératoires du carcinome hépatocellulaire. Celle-ci exploite des images TDM multi-phases et applique une forêt d’arbres décisionnels à des régions sémantiques par le biais de descripteurs dynamiques extraits à l’échelle des clusters. Comparée aux approches existantes, notre méthode permet une segmentation des tissus sains, actifs et nécrosés plus fiable et moins coûteuse en interaction opérateur.

Remerciement. Ce travail a reçu, pour sa réalisation, le soutien de la Fondation Arc : www.fondation-arc.org.

Références

- [1] Forner, A., Llovet, J.M., Bruix, J. : Hepatocellular carcinoma. *The Lancet* (2012)
- [2] Ronot, M., *et coll.* : Alternative response criteria (Choi, EASL and mRECIST) versus RECIST1.1 in patients with advanced hepatocellular carcinoma treated with *Sorafenib*. *The Oncologist* **19**(4) (2014) 394–402
- [3] Cantu, M., *et coll.* : Pathologic response to non-surgical locoregional therapies as potential selection criteria for liver transplantation for hepatocellular carcinoma. *Medical Science Monitor Basic Research* **18** (2013) 273–284
- [4] Shim, J.H., *et coll.* : Complete necrosis after transarterial chemoembolization could predict prolonged survival in patients with recurrent intrahepatic HCC after curative resection. *Annals of Surgical Oncology* **17**(3) (2010)
- [5] Lee, J., Cai, W., Singh, A., Yoshida, H. : Estimation of necrosis volumes in focal liver lesions based on multi-phase hepatic CT images. *Virtual Colonoscopy and Abdominal Imaging*. (2011) 60–67
- [6] Fang, R., Zabih, R., Raj, A., Chen, T. : Segmentation of liver tumor using efficient global optimal tree metrics graph cuts. *Abdominal Imaging. Computational and Clinical Applications* (2012) 51–59
- [7] Geremia, E., Menze, B., Clatz, O., Konukoglu, E., Criminisi, A., Ayache, N. : Spatial decision forests for MS lesion segmentation in multi-channel MRI. *Medical Image Computing and Computer-Assisted Intervention* (2010)
- [8] Cuingnet, R., Prevost, R., Lesage, D., Cohen, L.D., Mory, B., Ardon, R. : Automatic detection and segmentation of kidneys in 3D CT images using random forests. *Medical Image Computing and Computer-Assisted Intervention* (2012) 66–74
- [9] Breiman, L. : Random Forests. *Machine learning* **45**(1) (2001) 5–32
- [10] Avants, B.B., *et coll.* : Symmetric diffeomorphic image registration with cross-correlation : evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* **12**(1) (2008) 26–41
- [11] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S. : SLIC superpixels compared to SOA superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(11) (2012) 2274–2282
- [12] Warfield, S.K., Zou, K.H., Wells, W.M. : Simultaneous truth and performance level estimation (STAPLE) : an algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging* **23**(7) (2004) 903–921