2nd International Conference on Advanced Technologies
for Signal and Image Processing - ATSIP'2016
March 21-24, 2016, Monastir, Tunisia

Deep Random Forest-based Learning Transfer to SVM for Brain Tumor Segmentation

Samya AMIRI
SAGE laboratory
Higher Institute of Computer Science and Communication Techniques of Hammam sousse
Tunisia
Email: amiri.sam6@gmail.com

Islem REKIK
Department of Radiology and BRIC
University of North Carolina at Chapel Hill NC, USA
Email: islem.rekik@gmail.com

Mohamed Ali MAHJOUB
SAGE laboratory
National School of Engineers of Sousse
Tunisia
Email: medali.mahjoub@ipeim.rnu.tn

***Abstract*—Using neuroimaging techniques to diagnose brain tumors and detect both visible and invisible cancer cells infiltration boundaries motivated the emergence of diverse tumor segmentation algorithms. Noting the large variability in both tumor appearance and shape, the task of automatic segmentation becomes more difficult. In this paper, we propose a random-forest (RF) based learning transfer to SVM classifier method for segmenting tumor lesions while capturing their complex characteristics. Our framework is composed of two cascaded stages. In the first stage, we train a random forest to learn the mapping from the image space to the tumor label space. In the testing stage, we use the predicted label output from the random forest and feed it along with the testing intensity image to an SVM classifier to get the refined segmentation. Then we make our RF-SVM cascaded classification steps deep through an iterative process. We tested our method on 20 patients with high-grade gliomas from the Brain Tumor Image Segmentation Challenge (BRATS) dataset. Our proposed framework significantly outperformed SVM-based segmentation and RF-based segmentation –when used solely.**

***Keywords*—Segmentation, Brain tumor, Random Forest, SVM, MRI**

I. Introduction

High-grade glioma is the most frequent primary brain cancer. It represents 85% of new cases of malignant primary tumor diagnosed every year. Magnetic resonance imaging (MRI) plays an important role in brain glioma diagnosis and treatment planning strategies [1]. Particularly, accurately outlining the tumor visible infiltration boundary can help quantify the lesion size and its progression rate in longitudinal data. However, noting the anisotropic complex shape that high-grade gliomas can take due to their high diffusion and proliferation rates, as well as their appearance variability, the segmentation of the complete glioma including all its compartments (e.g. necrotic core, edema) becomes more challenging [2]. As manual segmentation is time consuming and is prone to inter-observer variability, automatic segmentation is desired.

Medical imaging segmentation is widely addressed in the state-of-the art as a classification problem where the proposed methods can be divided into two main classes. The first class includes discriminative segmentation methods, which are mainly based on image features and the training data. The second class includes generative approaches that require additional information about the space domain.

Recently, random forest-based segmentation methods have shown good performances and robustness. Noting the lack of extensive use of random forest for high-grade gliomas segmentation and since stroke lesions are as challenging in segmentation as brain gliomas, many recent papers also used random forest for lesion segmentation. For instance, in [3], standard random forests of 100 trees were used to independently classify pixels belonging to stroke lesions. The used features were the intensities associated with each pixel. Post-processing operations were performed to smooth the probabilities for each slice between different lesion parts.

Based on the implementation of the random forest presented in [4], the authors of [5] proposed a stroke lesion segmentation method. Each voxel was associated with many features including: intensity, Gaussian, hemispheric difference, local histogram and center distance. The mean Dice score reached about 0.58 with a standard deviation of 0.29. Halme et al. proposed a three-level segmentation approach in [16], where they combined prior segmentation with a feature set to perform a random forest-based segmentation. Then, they refined the segmentation results using contextual clustering based on Markov random field and iterated conditional modes. In [7], after correction and normalization, a set of features was extracted from the MRI sequence: intensities, smooth intensities (A Gaussian), median intensities, gradient, magnitude of the gradient and local entropy. These features were used to train a random forest of 150 tree. The mean Dice score reached 0.54, with a standard deviation of 0.26.

Besides the classification techniques, the use of the appropriate set of feature is of great importance and affects the segmentation results [19]. In this context several approaches are based on the simultaneous use of local and context sensitive features, which greatly enhances the segmentation results. In [16], the authors propose an iterative two-stage probabilistic graphical model composed of two levels of Markov Random Fields (MRFs). At the bottom level, potential lesion voxels are identified using local and neighbourhood intensities and class priors. Regions are constructed according to the tissue type. The higher level, evaluates the probability of candidate lesions, based on group intensity, texture and neighbouring regions. The inference of each level is propagated to the other one iteratively to suppress false positives and refine lesion boundaries. The Dice Index For this method is 0.69, its strength relies on detecting small lesions.

978-1-4673-8526-8/16/$31.00 ©2016 IEEE

In [17], [18], the authors propose several architectures based on a combination of modified versions of the Convolutional Neural Network (CNN) to exploit both local and large contextual features. These methods are applied on 3D data.

In this paper, we present a fully automatic segmentation method based on a standard random forest (RF) algorithm and SVM. More specifically, we propose to use a random-forest based learning transfer to SVM classifier method for segmenting high-grade gliomas. Our framework is composed of two cascaded stages. In the first stage, we train a random forest using flair MRI image intensity features to learn the mapping from the image space to the tumor label space. In the testing stage, we use the predicted label output from the random forest and pass it along with the testing intensity image to an SVM classifier to get the refined segmentation. For more refinement, we deepen the classification process by re-iterating these two steps and using the outcome of one RF-SVM cascade into the next cascade. We evaluated our proposed framework on 20 high-grade glioma patients from the Brain Tumor Image Segmentation Challenge (BRATS[1]) dataset [13]. Our method significantly outperformed both SVM-based and RF-based segmentation when used solely.

II. Deep random forest-based learning transfer to SVM for segmentation

A. Random Forests

Ensemble learning techniques are based on the aggregation of the results generated by multiple weak predictors or learners. Their main principle is that many predictors (i.e. trees) give a larger exploration of the solution space, then the resulting predictor is more powerful than each individual one [8]. Among these ensemble techniques, random forest showed promising results in several applications. Firstly, introduced by Beirman [9], RF represents a family of methods based on the aggregation of decisions output by many tree-like weak learners. This family can be divided into two main classes: Classical Random Forests and Fully Random Forests. In the first class, training data samples are used to partition the input space; while in the second class, a fully random partition is generated independently of the training data samples that will be used for labeling the final partition [10].

A random forest independently creates a tree for each bootstrap sample generated from the data Figure 1. Unlike standard trees, the RF tree construction is as follows. To split a node, a number m of features is randomly picked, then we split according to the best feature among the m selected features. In addition, the built tree is completely developed (maximal tree) and is not pruned. Thus, Random Forests are usually considered as a variant of Bagging (randomized bagging), where the main difference resides in the construction of the trees, making the aggregation trees more independent [8], [10]. Random Forests are suitable for several applications in different domains because they are fast, resistant to over-fitting and have a good performance in high-dimensional data [11], [14].

[1] http://braintumorsegmentation.org/

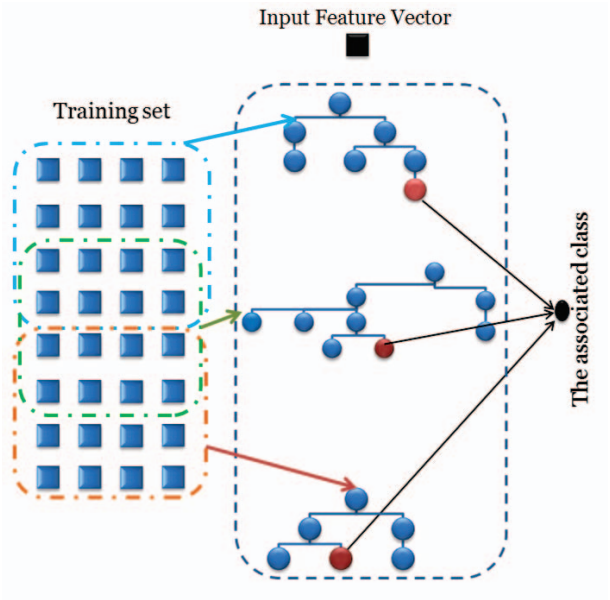


Fig. 1: *Random forest architecture: a set of trees is built using random subsets of the training data. For an input data the classification is performed by each tree according to its feature vector. Red points represent the individual classification result; the associated class is decided by major vote.*

B. Linear Support Vector Machine

Support vector machine is a supervised learning technique firstly introduced by Vladimir Vapinik [12], and which demonstrated remarkable performances especially for data analysis and pattern recognition. SVM belongs to the class of regularization learning machines that perform linear learning in non-linear spaces. The main idea of SVM is to reformulate the classification problem as a quadratic programming optimization. Hence, data classification consists in finding the best separating hyperplane that maximizes the margin (the distance to the nearest data point of the classes, see Figure 2).

In the case of nonlinear classification, a kernel function (Polynomial, Gaussian Radial Basis Function, Multi-Layer Perceptron) is used to transform the feature space to a Higher-dimension space where a linear separation is possible [15]. Given a training data set , to find the maximum-margin hyperplane new data points are classified according to their distance from the hyperplane.

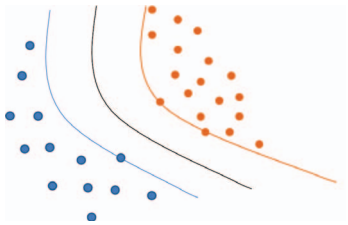


Fig. 2: *Two classes Nonlinear SVM: The data are represented by blue and orange dots. Maximum-margin hyperplane (in black) is the hyperplane with the largest distance to the nearest data of each class.*

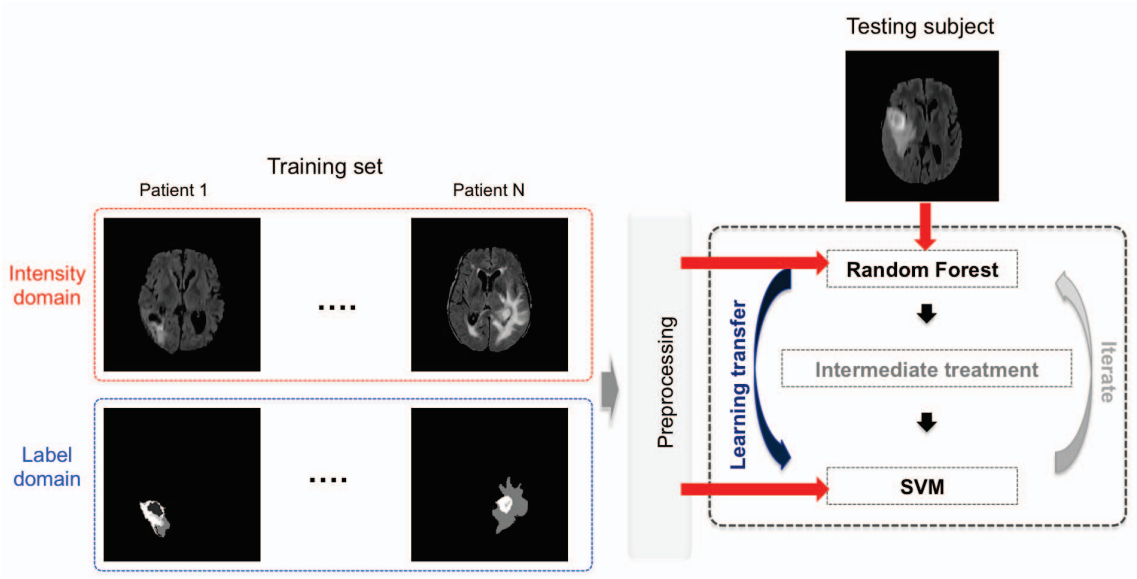


Fig. 3 *Flowchart of the proposed method: the training set is composed of brain MRIs and their corresponding tumor segmentation images, after preprocessing step, the extracted features along with the labels' set are used to train both Random Forest and SVM. For an input subject the two classifiers (RF and SVM) communicate iteratively the segmentation result for refinement. An intermediate treatment is performed to correct RF segmentation result.*

C. Deep RF-learning transfer to SVM for glioma segmentation

In this paper, we propose a brain tumor segmentation algorithm based on a combination of standard random forest and SVM. The architecture of the proposed framework is composed of two major steps as shown in Figure 3: first preprocessing and feature extraction; second, training RF and SVM; and finally, testing and generating the final segmentation results.

In the first stage, we train a random forest to learn the mapping from the image space to the tumor label space. In the testing stage, we use the predicted label output from the random forest and feed it along with the testing intensity image to an SVM classifier to get the refined segmentation. Then we make our RF-SVM cascaded classification steps deep through an iterative process.

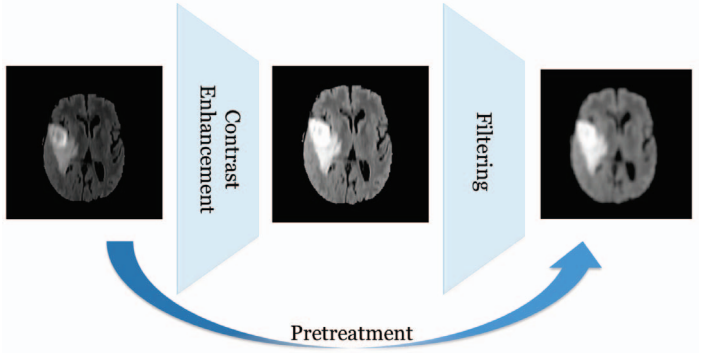


Fig. 4 *Image preprocessing steps. For each patient, we first enhance the input image and then we perform and filtering and normalization.*

To generate meaningful and discriminative features for RF training, we add a few preprocessing steps.

This includes filtering, contrast enhancement and intensity normalization (Figure 4). We then use first order features (intensity, mean and median) to characterize each pixel. During the training stage, we use these features to learn a non-linear mapping between the input features and the labels by training a random forest.

Next, in the testing stage, we use the aggregated RF prediction label map and the same features to independently train an SVM classifier.

Given a preprocessed input 2D image, an iterative classification process is applied. First, Random Forest classifies the foreground pixels and generates a prior segmentation that will be forwarded to the SVM classifier. The latter, will consider a large region of interest (ROI) centralized on the prior segmentation. SVM-based classification will be performed on this ROI except the prior segmentation.

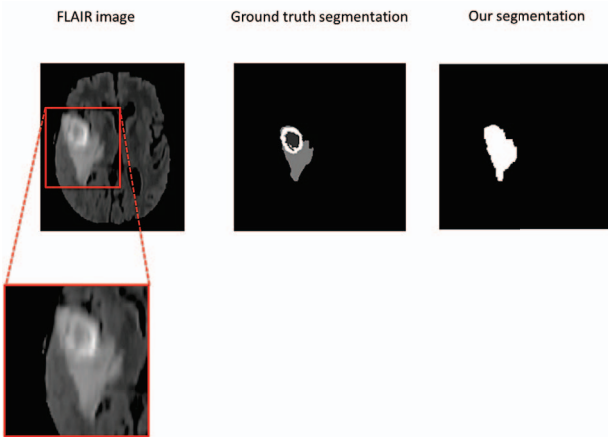
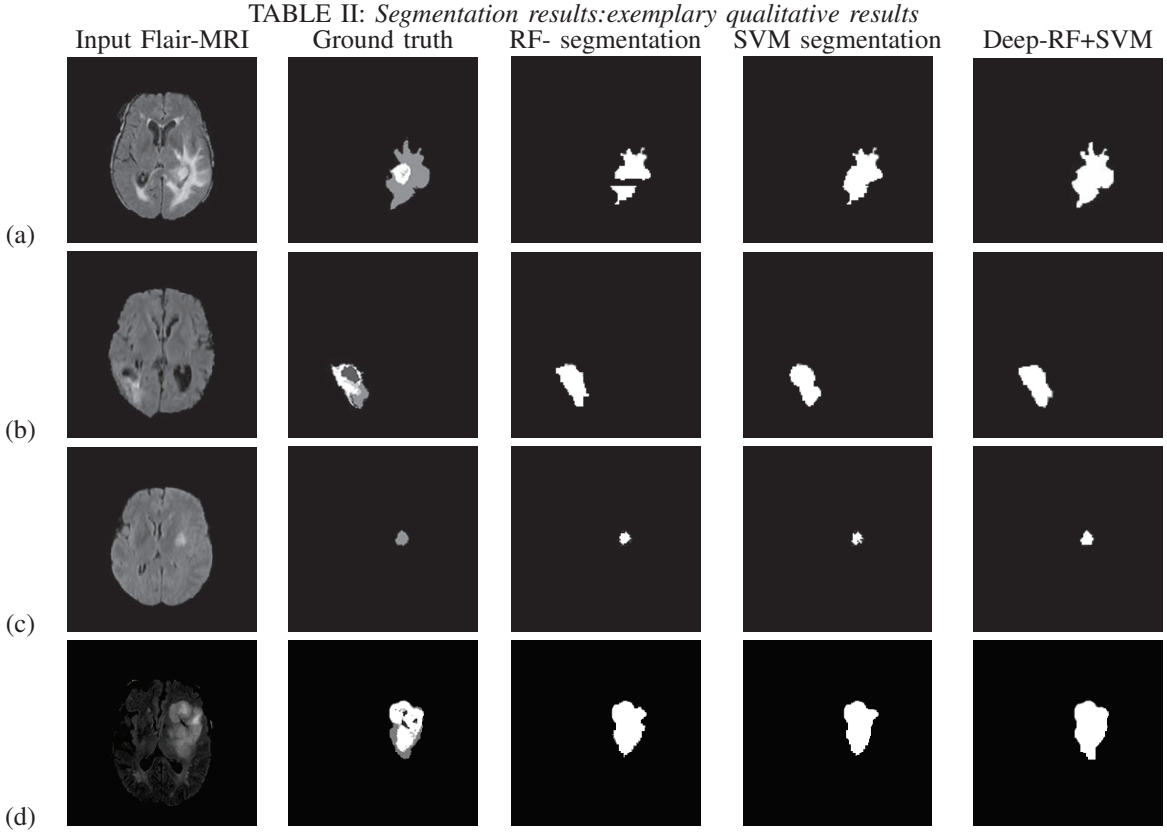


Fig. 5 *An example of segmentation using our deep RF-learning transfer to SVM.*

TABLE II: *Segmentation results:exemplary qualitative results*



Hence, SVM explores the neighborhood of the resulting RF segmentation. Then, we will recall the random forest to reclassify the indicated ROI. A threshold of the mis-classified region size (number of pixels) is defined.

These two steps will be iterated to further refine the segmentation results. Of note, since the segmentation result of the random forest may contain some holes (black pixels) representing the mis-classified pixels, we first fill in these holes before passing the RF-label map to SVM, Figure 5 shows An example of segmentation using our deep RF-learning transfer to SVM.

The strength of the proposed framework relies on iterating classification only for doubtful zones. In addition, the classification by standard RF is based only on local features while contextual information is handled by SVM; thus the learning transfer is not only a simple communication between two classifiers, it is a combination of both local and contextual features. This idea is basically used/shared by several research works [16]–[18].

III. Experiments

In this section, we will present an experimental analyze of the proposed method. Experimental settings, dataset and results are described.

A. Dataset

The Brain Tumor Image Segmentation Challenge (BRATS) dataset [13] contains about 300 cases of both low-grade and high-grade glioma.

For each case, T1 MRI, T1 contrast-enhanced MRI, T2 MRI, T2 FLAIR MRI volumes are available. Each subject's image volumes are rigidly co-registered to the T1c MRI, and resampled to a common resolution while glioma is manually labeled.

For our experiments, we selected 20 random patients with high-grade gliomas from the Brain Tumor Image Segmentation Challenge (BRATS) dataset. Using this dataset allows to compare our approach to state-of-the-art of tumor segmentation. Only FLAIR MR images are used for testing and training the proposed framework.

B. Parameters setting

Our method has been implemented in Matlab. For random forest, the number of trees T is fixed to 100 while the depth of each tree is 50. The number of iterations of learning transfer between SVM and RF is N=3, the threshold size *Siz* is the number of the mis-classified pixels according to the SVM stage. It is fixed to 20 pixels. The computational time is related to the chosen parameters, the size of the image and the training set.

N and Siz are empirically fixed such as we have good segmentation within an acceptable computational time. The size of the region judged to be misclassified by SVM indicates the correspondence and the harmony between the SVM and RF classification. We studied the influence of the iteration number on the region size, as shown in Figure 6 they are inversely proportional, thus when the number if iterations increases, the number of the mis-classified pixels by SVM decreases.

TABLE I: *Segmentation results evaluation using mean Dice index, averaged across 20 patients.*

Method	SVM	RF	RF + SVM
Mean Dice index	0.59 ± 0.10	0.63 ± 0.09	$\mathbf{0.72 \pm 0.11}$
paired t-test (p-value)	$p = 0.0002$	$p = 0.004$	–

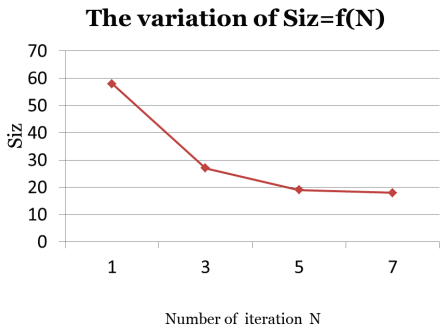


Fig. 6 *Number of mis-classified pixels variation in function of iteration number (N).*

C. Segmentation results

Classifiers were trained using leave-one-patient-out experiments. The quality of the obtained segmentation of the left-out patient was evaluated on the basis of the manually annotated ground truth using the well-known DICE index between two regions R1 and R2 computed as:

$$D\left(R1,R2\right)=\frac{2|R1\bigcap R2|}{|R1|+|R2|}.$$

The segmentation results of our method are compared to those of RF and SVM separately. Table I shows the mean DICE index over the 20 patients for each method. The first column shows the segmentation results when solely using SVM, the second column when using RF and the third column our proposed framework.

Our method clearly outperforms both SVM and RF with a high statistical significance using paired t-test. According to mean DICE index, the proposed method is obviously improving the segmentation results with mean DICE index 72% compared to SVM (59%) and RF (63%).

TableII provides some exemplary qualitative results. The effectiveness of our method is notable for large tumors (tableII (a)), while the performances for small tumors of the three tested methods (RF, SVM and RF+SVM) are comparable (tableII (c)). In addition, our method gives better results for fine details of the tumor boundary (tableII(b)).

The proposed framework is based on a cascaded classification of RF and SVM. A deep learning transfer is performed between the two classifiers to exploit both local features as well as contextual features simultaneously. The results prove the efficiency of the concept and can be surely optimized by using a rich feature vector or optimal versions of RF or SVM. Thus, compared to other methods based on this concept [16] The proposed deep RF-learning transfer to SVM approach is simple and fast (average computation time of training and testing is about 2 min) with comparable Dice index.

IV. Conclusion

In this paper, we presented a simple and reliable brain tumor segmentation method in MRI images through recursively and deeply transferring a learnt random forest-based knowledge to guide an SVM classifier in the segmentation task. The proposed method has an acceptable performances compared to other approaches in the literature [13]. Most importantly, it is simple and fast and it also significantly ($p < 0.005$) outperformed the mere use of SVM or RF for tumor classification. As our method is only based on first order features, we aim to refine the developed framework by including contextual features. Moreover, we will investigate the performance of the random forest when combined with other robust classifiers such as neural networks.

Acknowledgment

The authors gratefully thank BRATS research team for kindly sharing their data.

References

[1] Narayanan, V. *High Grade Gliomas: Pathogenesis*, Management.

[2] I Despotovi, B Goossens and W Philips, *MRI Segmentation of the Human Brain: Challenges, Methods, and Applications*. Computational and mathematical methods in medicine, 2015.

[3] O Maier, M Wilms and H Handels, *Random forests with selected features for stroke lesion segmentation*. Ischemic Stroke Lesion Segmentation, page 17 2015.

[4] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. *Scikit-learn: Machine learning in Python*. Journal of Machine Learning Research, 12:28252830, 2011.

[5] L. Chen, P. Bentley, D. Rueckert, *A novel framework for sub-acute stroke lesion segmentation based on random forest*. Ischemic Stroke Lesion Segmentation, page 9, 2015.

[6] H.L. Halme, A. Korvenoja, E. Salli, *segmentation of stroke lesions using spatial normalization*, random forest classification and contextual clustering. Ischemic Stroke Lesion Segmentation, page 29, 2015.

[7] M. Qaiser and A. Basit. *Automatic ischemic stroke lesion segmentation in multi-spectral mri images using random forests classifier*. Ischemic Stroke Lesion Segmentation, page 43, 2015.

[8] A. Liaw, M. Wiener, (2002). *Classification and regression by randomForest*. R news, 2(3), 18-22.

[9] L. Breiman, *Random forests*. Machine learning, 2001, 45(1), 5-32.

[10] R. Genuer, (2010). *Forets aleatoires: aspects theoriques, selection de variables et applications*. (Doctoral dissertation, Universit Paris Sud-Paris XI).

[11] H. Chen, Z. Lin, H. Wu, L. Wang, T. Wu and C. Tan *Diagnosis of colorectal cancer by near-infrared optical fiber spectroscopy and random forest*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2015, 135, 185-191.

[12] C. Cortes and V. Vapnik, *Support-vector networks*. Machine learning, 1995, 20(3), 273-297.

[13] B. Menze, M. Reyes, K. Van Leemput, *The Multimodal Brain TumorImage Segmentation Benchmark (BRATS)*. IEEE, 2014.

[14] H. Shili, L. B. Romdhane, and B. Ayeb, *Reliable Probabilistic Classification of Mammographic Masses using Random Forests*. In The 9th International Conference on Data Mining)(Las Vegas, Nevada, USA), 2013.

[15] C. J. Burges, *A tutorial on support vector machines for pattern recognition*. Data mining and knowledge discovery, 2(2), 121-167, 1998.

[16] N. Subbanna, D. Precup, D. Arnold and T. Arbel, *Image: Iterative multilevel probabilistic graphical model for detection and segmentation of multiple sclerosis lesions in brain mri*. In Information Processing in Medical Imaging, pages 514526. Springer, 2015.

[17] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y.Bengio, Chris Pal, P.M. Jodoin and H. Larochelle, *Brain tumor segmentation with deep neural networks*. arXiv preprint arXiv:1505.03540, 2015.

[18] K. Kamnitsas, L. Chen, C. Ledig, D. Rueckert and B. Glocker. *Multi-scale 3d convolutional neural networks for lesion segmentation in brain mri*. Ischemic Stroke Lesion Segmentation, page 13, 2015.

[19] R. Meier, S. Bauer, J. Slotboom, R. Wiest and M. Reyes. *Appearance- and context-sensitive features for brain tumor segmentation*. MICCAI BRATS Challenge Proceedings (2014).