Two stages CNN-based segmentation of gliomas, uncertainty quantification and prediction of overall patient survival

Thibault Buatois, Élodie Puybareau, Guillaume Tochon, and Joseph Chazalon

EPITA Research and Development Laboratory (LRDE), France
surname.name@lrde.epita.fr

Abstract. This paper proposes, in the context of brain tumor study, a fast automatic method that segments tumors and predicts patient overall survival. The segmentation stage is implemented using two fully convolutional networks based on VGG-16, pre-trained on ImageNet for natural image classification, and fine tuned with the training dataset of the MICCAI 2019 BraTS Challenge. The first network yields to a binaty segmentation (background vs lesion) and the second one focuses on the enhancing and non-enhancing tumor classes. The final multiclass segmentation is a fusion of the results of these two networks. The prediction stage is implemented using kernel principal component analysis and random forest classifiers. It only requires a predicted segmentation of the tumor and a homemade atlas. Its simplicity allows to train it with very few examples and it can be used after any segmentation process.

Keywords: glioma · tumor segmentation · fully convolutional network · random forest · survival prediction.

1 Introduction

1.1 Motivation

Gliomas are the most common brain tumors in adults, growing from glial cells and invading the surrounding tissues [10]. Two classes of tumors are observed. The patients with the more aggressive ones, classified as high-grade gliomas (HGG), have a median overall survival of two years or less and imply immediate treatment [13, 16]. The less aggressive ones, the low-grade gliomas (LGG), allow an overall survival of several years, with no need of immediate treatment. Multimodal magnetic resonance imaging (MRI) helps pratitioners to evaluate the degree of the disease, its evolution and the response to treatment. Images are analyzed based on qualitative or quantitative measures of the lesion [8, 22]. Developing automated brain tumor segmentation techniques that are able to analyze these tumors is challenging, because of the highly heterogeneous appearance and shapes of these lesions. Manual segmentations by experts can also be a challenging task, as they show significant variations in some cases. Despite

the relevance of glioma segmentation, this segmentation is challenging due to the high heterogeneity of tumors. The development of an algorithm that can perform fully automatic glioma segmentation and overall prediction of survival would be an important improvement for patients and practitioners. During the past 20 years, different algorithms for segmentation of tumor structures has been developed and reviewed [1, 6, 7]. However, a fair comparison of algorithms implies a benchmark based on the same dataset, as it has been proposed during MICCAI BraTS Challenges [15] [5].

1.2 Context

The work proposed in this article has been done in the context of the MICCAI 2019 Multimodal Brain Tumor Segmentation Challenge (BraTS)[1]. The overall goal of this challenge is to establish a fair comparison between state-of-the-art methods, and to release a large annotated dataset. The objective of the work conducted within the scope of BraTS challenge is three-fold:

Task 1 providing a fully automated pipeline for the segmentation of the glioma from multimodal MRI scans without any manual assistance,

Task 2 predicting the patient overall survival from pre-operative scans,

Task 3 estimating the uncertainty in segmentation results provided within the scope of Task 1. Note that, unlike the two previously mentioned tasks which were already established in the former BraTS challenges, Task 3 is exclusive to the 2019 BraTS challenge.

We received data of 335 patients, with associated masks to develop our method. The data, avalaible online, have been annotated and preprocessed [2–4]. The volumes given are T1, T1ce, T2 and FLAIR. Our method is then evaluated on new volumes:

- a validation set composed of 125 patients, released by the organizers without the manual segmented masks (these masks will not be released), to obtain preliminary results.
- a test set comprising 166 patients, used for the challenge evaluation. In that case, our team was asked to process those data and send the corresponding results in the expected formats within 48h after receiving the data.

1.3 Related works

In the framework of BraTS 2018 challenge, we carried out the tumor segmentation task thanks to a fully convolutional network (FCN) approach [17]. The used network was VGG (Visual Geometry Group) [20], pre-trained on the ImageNet dataset and fine-tuned to the BraTS challenge dataset thanks to transfer learning strategy. As VGG expects 2D color images as input images (thus, 2D

[1] `http://braintumorsegmentation.org/`

images with 3 RGB channels), we made use of the "pseudo-3D" strategy, originating from [23]. Initially applied to the segmentation of 3D brain MR volumes, the pseudo-3D idea consists in creating a series of RGB images of the 3D input volume by selecting the $(n-1)^{\text{th}}, n^{\text{th}}$ and $(n+1)^{\text{th}}$ slices (with n sliding through the 3D volume) as the R, G and B channels, respectively. This pseudo-3D idea proved to be also successful for left atrium segmentation [18]. Note that in [17], this pseudo-3D idea was actually used in a multimodality fashion since the $(n-1)^{\text{th}}$ and $(n+1)^{\text{th}}$ slices were extracted from the T1ce modality while the n^{th} slice was selected in the T2 modality.
However, this method yielded poor results for the tumor segmentation task of BraTS 2018 challenge. Thus, we decided to improve our approach, still using VGG and 3D-like images, but with a two-stage segmentation scheme, as detailed in the following section 2.1. Based on those poor quality tumor segmentations, we nevertheless reached the second place for the survival prediction task of BraTS 2018 challenge [17]. We also attempt to improve our survival prediction algorithm, as exposed in section 2.2.

2 Proposed methodology

2.1 Tumor segmentation

Overall FCN architecture As for the BraTS 2018 Challenge [17], our FCN architecture for the segmentation task relies on the 16-layer VGG network [20], which was pre-trained on ImageNet for image classification purposes [11]. We keep only the first 4 convolutional stages, and we discard the fully connected layers at the end of VGG network. Each stage is composed of convolutional layers, followed by Rectified Linear Unit (ReLU) layers and a max-pooling layer, correponding to four fine-to-coarse feature maps. Inspired by the work in [12,14], we add specialized convolutional layers (with a 3×3 kernel size) with K (e.g. $K = 16$) feature maps after the last convolutional layer of each stage. All these specialized layers are then rescaled to the original image size, and concatenated together. We add a last convolutional layer with kernel size 1×1 at the end. This last step combines linearly the fine-to-coarse feature maps in the concatenated specialized layers, and provide the final segmentation result. The parameters of the network are the same than in [17].

This FCN architecture is depicted in its generic form in Fig 1, where it accepts as input a RGB image and outputs its corresponding segmentation map.

Pre-processing Let n, m be respectively the minimum non-null and maximum gray-level value of an input 3D volume. For each patient, we first requantize all voxel values using a linear function so that the gray-level range $[n, m]$ is mapped to $[-127, 127]$.

For our application, the question amounts to how to prepare appropriate inputs (RGB input images) given that a brain MR image is a 3D volume. To

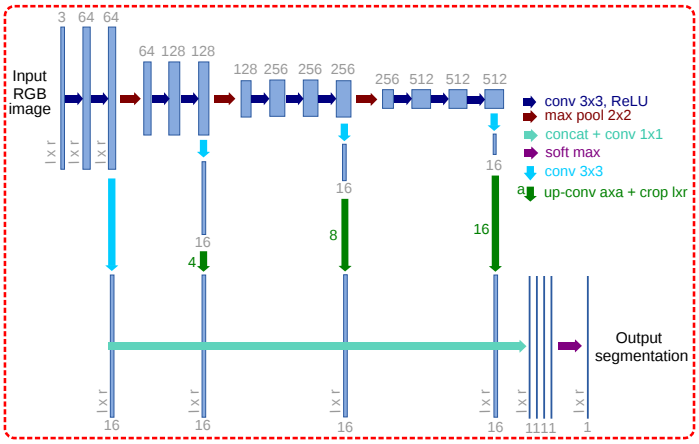

Fig. 1. Proposed FCN generic architecture, based on the pre-trained VGG network.

that aim, we propose to stack 2D slices of different modalities according to the segmentation classes studied.

Precisely, to form an input artificial color image for the pre-trained network to segment the i^{th} slice, we defined the input image this way:

- in the channel green, put the slice i of modality 1
- in the channel red, put the slice $n - x$ of modality 2
- in the channel blue, put the slice $n + x$ of the modality 3

The x parameter, namely the "offset", can bring 3D information for 2D segmentation while the choice of modalities (modality 1, 2 and 3 can be the same) can bring information from one to three modalities. For the challenge, we use the slice i of one volume modality in the green channel, the slice i of an other modality for the blue and red channels, and the offset x is set to 0. These parameters (modalities and x) have been selected after testing all the combinations.

Segmentation During our experiments, we noticed that our network confuses the background with the edema, and the necrotic and non-enhancing tumor core with the enhancing tumor. Hence, we decided to performed the segmentation in two steps: the first step is a binary segmentation of the whole tumor (vs. background), and the second is a multi-label segmentation of two parts of the tumor (necrotic and non-enhancing tumor core vs. enhancing tumor). The final segmentation resulted of the fusion of these two results, as depicted in Fig. 2.

As the different modalities enhance different structures, we decided to take the advantage of each modality for a specific task. For the segmentation of the whole tumor vs. background, we put the T2 modality in the green channel of our RGB input image and the FLAIR in the blue and red channels. For the segmentation of necrotic and non-enhancing tumor core vs. enhancing tumor,

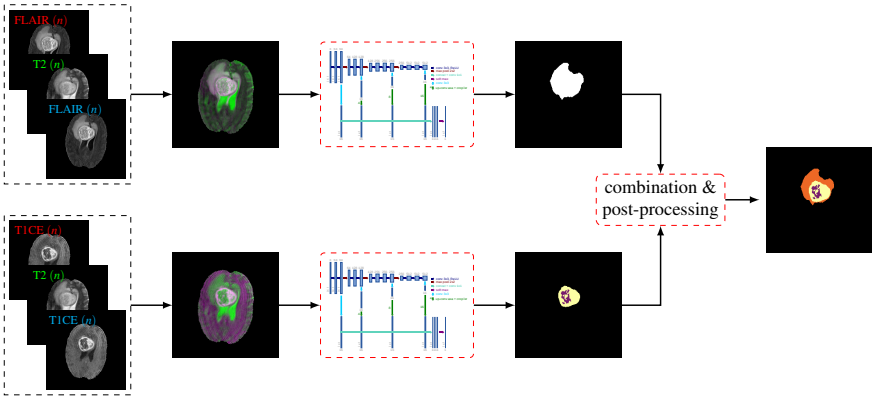

Fig. 2. Workflow of our procedure for the segmentation task.

we put the T2 modality in the green channel of our RGB input image and the T1ce in the blue and red channels.

Post-processing The output of our process for one slice during the inference phase is two 2D segmented slice: one for the tumor vs. background, and one for the inner parts of the tumor. We first fused the two segmentation results: we kept the segmentation of the inner parts of the tumor as classes *necrotic and non-enhancing tumor core* and *enhancing tumor*. The last class, the *oedema* is the non-zero part of the whole tumor that was not detected as an other class in the tumor part segmentations.

After treating all the slices of the volume, the segmented slices are stacked to recover a 3D volume with the same shape as the initial volume, and containing only the segmented lesions. A spatial regularization is processed to ensure 3D cohesion, and we remove the 3D connected components that have as label only *oedema*.

2.2 Patient survival prediction

The second task of the MICCAI 2019 BraTS challenge is concerned with the prediction of patient overall survival from pre-operative scans (only for subjects with gross total resection (GTR) status). As precognized by the evaluation framework, the classification procedure is conducted by labeling subjects into three classes: short-survivors (less than 10 months), mid-survivors (between 10 and 15 months) and long-survivors (greater than 15 months). For post-challenge analyses, prediction results are also compared in terms of mean and median square error of survival time predictions, expressed in days. For that reason, our proposed patient survival prediction algorithm is organized in two steps:

1 We first predict the overall survival class, *i.e.* short-, mid- or long-survival (hereafter denoted by class/label $1, 2$ and 3, respectively).
2 We then adjust our prediction within the predicted class by means of linear regression, in order to express the survival time in days.

Definition and extraction of relevant features Extracting relevant features is critical for classification purposes. Here, we re-use the features implemented by our team in the framework of the patient survival prediction task of MICCAI 2018 BraTS challenge, which ranked tie second [17]. Those features were chosen after in-depth discussions with a practicioner and are the following:

feature 1: the patient age (expressed in years).
feature 2: the relative size of the necrosis (labeled 1 in the groundtruth) class with respect to the brain size.
feature 3: the relative size of the edema class (labeled 2 in the groundtruth) with respect to the brain size.
feature 4: the relative size of the active tumor class (labeled 4 in the groundtruth) with respect to the brain size.
feature 5: the normalized coordinates of the binarized enhanced tumor (thus only considering necrosis and active tumor classes).
feature 6: the normalized coordinates of the region that is the most affected by necrosis, in a home made brain atlas.

For the training stage, features $2, 3$ and 4 are computed thanks to the patient ground truth map for each patient. As this information is unknown during the test stage, the segmented volumes predicted by our Deep FCN architecture are used instead. In any case, these size features are expressed relatively to the total brain size (computed as the number of voxels in the T2 modality whose intensity is greater than 0).
In addition, we also re-use the home-made brain atlas that we also developed for the 2018 BraTS challenge. This atlas is divided into 10 crudely designed regions accounting for the frontal, parietal, temporal and occipital lobes and the cerebellum for each hemisphere (see [17] for more details regarding this atlas and how it is adjusted to each patient brain size). Feature 6 is defined as the coordinates of the centroid of the region within the altas that is the most affected by the necrosis class (*i.e.*, the region that has the most voxels labeled as necrosis with respect to its own size). Note that this feature, as well as feature 5, is then normalized relatively to the brain bounding box. This leads to a feature vector with 10 components per patient (since both centroids coordinates are 3-dimensionals).

Training phase For the training phase, we modified our previous work [17] in the following way: while we maintained the final learning stage through random forest (RF) classifiers [21], we replaced the principal component analysis (PCA) transformation, acting as preprocessing step for the learning stage, by its kernel counterpart (kPCA) [19]. The rationale is that we hope to increase the RFs

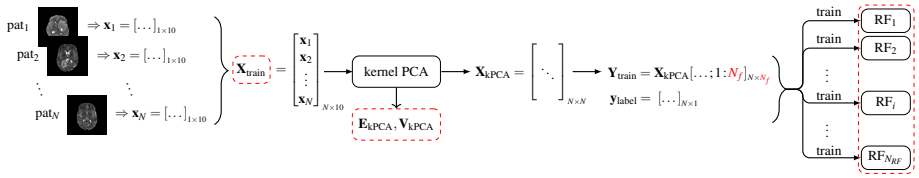

Fig. 3. Workflow of the proposed class-based training procedure. The information stored after the training phase (necessary for the test phase) is written in red or encircled in dashed red.

performances in terms of classification/prediction as the input features are highly non-linear in terms of survival labels.
More specifically, the training stage of our prediction algorithm is as follows:

1. The feature vector $\mathbf{x}_i \in \mathbb{R}^{10}$ of each of the N patients in the training set is extracted as described in the previous section 2.2. All those feature vectors are then stacked in a $N \times 10$ feature matrix $\mathbf{X}_{\text{train}}$
2. A kPCA is performed on $\mathbf{X}_{\text{train}}$, yielding the $N \times N$ matrix $\mathbf{X}_{\text{kPCA}}$. This matrix is obtained through the computation, normalization and diagonalization of the so-called *kernel matrix* which represents the dot product between the N features vectors when mapped in the feature space through a kernel function (here defined as a polynomial kernel with degree $d = 3$).
4. The $N \times N_f$ matrix $\mathbf{Y}_{\text{train}}$ is defined from $\mathbf{X}_{\text{train}}$ by retaining the first N_f columns (corresponding to the leading N_f features in the feature space, here set to $N_f = 10$). N_{RF} RF classifiers [21] are finally trained on all rows of $\mathbf{Y}_{\text{train}}$ to learn to predict the survival class of each training patient using the true label vector $\mathbf{y}_{\text{label}}$ as target values. The used RF parameters (number of decision trees per RF, splitting criterion, total number of RFs N_{RF}) are defined as in [17].
5. Three linear regressors (one per survival class) are finally trained using the patient age and its whole tumor size (relatively to its brain size) as explanatory variables and its true survival time (expressed in days) as measured variable.

Steps 1. to 4. are depicted by the workflow in Fig.3. In addition to the three linear regressors, we also store (for the test phase) the training feature matrix $\mathbf{X}_{\text{train}}$, the eigenvector matrix V_{kPCA} and eigenvalues E_{kPCA} of the kernel matrix, and the number of retained features N_f after kPCA.

Test phase The test phase is conducted in a similar fashion as the training phase. Given some input test patient, its overall survival class is first predicted, before being refined and expressed in terms of number of days. More specifically:

1. The features vector $\mathbf{x}_{\text{test}}$ of the test patient is retrieved as described previously.

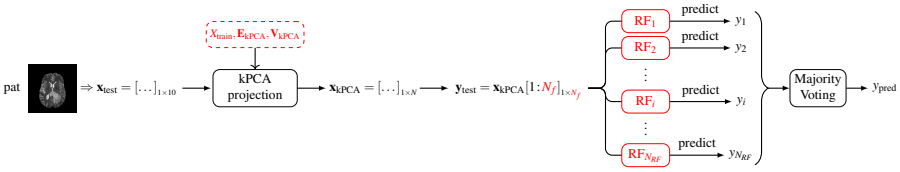

Fig. 4. Workflow of the proposed test procedure.

2. This feature vector is then projected onto the principal axes learnt by the kPCA during the training phase. For that purpose, a new kernel matrix is computed and centered (hence the need for $\mathbf{X}_{\text{train}}$) before proper projection (through $\mathbf{V}_{\text{kPCA}}$) and scaling (with $\mathbf{E}_{\text{kPCA}}$).
3. This results in the projected vector $\mathbf{x}_{\text{kPCA}} \in \mathbb{R}^N$ from which the first N_f features are retained, yielding the test vector $\mathbf{y}_{\text{test}}$. This vector is then fed to the N_{RF} RF classifiers, leading to N_{RF} independent class label predictions. The final label prediction y_{pred} (1, 2 and 3 for short-, mid- and long-survivors, respectively) is eventually obtained by majority voting.
4. Once the survival class has been established, the final patient survival rate is predicted by means of the appropriate learnt linear regressor.

Steps 1. to 3. are illustrated by the worflow in Fig.4.

2.3 Quantification of uncertainty in segmentation

The third task of MICCAI 2019 BraTS challenge focuses on the estimation of the uncertainty in segmentation results provided within the scope of Task 1. For that purpose, we focused on the study of a lightweight technique which estimate uncertainty by considering the instability at the spatial boundary between two regions predicted to belong to different classes. We believe that such approach can be complementary to approaches based on the stability of the prediction under perturbations like Monte Carlo Dropout [9] which tend to be computationally demanding.

The resulting indicator assigns a maximal uncertainty (100) at the boundary between two regions, and linearly decreases this uncertainty to the minimal value (0) at a given distance from the boundary. This distance defines the (half) width of an "uncertainty border" between two regions. It is calibrated independently for each class, and was estimated according to the 95^{th} percentile of the Hausdorff distance metric reported in Table 1 for our segmentation method for this particular class. In practice, we used a half-width of 9 voxels for the whole tumor (WT), a half-width of 12 voxels for the tumor core (TC) and a half-width of 7 voxels for the enhancing tumor (ET).

To compute this indicator, we first compute the Boundary Distance Transform $\text{BDT} = max(\text{DT}(\mathcal{R}), \text{DT}(\overline{\mathcal{R}}))$ using the Distance Transform DT to the given sub-region $\mathcal{R}$ and its complement $\overline{\mathcal{R}}$. Then, we invert, shift and clip the BDT such that the map is maximal on the boundary and have 0 values at a distance

Table 1. Dice and Haussdorf distance metrics for the proposed segmentation method on the test data set.

Metric	Dice_ET	Dice_WT	Dice_TC	HD95_ET	HD95_WT	HD95_TC
Mean	0.750	0.854	0.800	3.084	7.047	5.961
StdDev	0.230	0.129	0.247	4.320	7.609	8.847
Median	0.821	0.897	0.891	2	5.099	3.606
25quantile	0.733	0.838	0.821	1.414	3.742	2
75quantile	0.877	0.919	0.936	3.162	7.071	6.224

Table 2. Classification metrics of the proposed survival prediction method for the validation (with comparison with [17]) and test data sets.

	Data set	Accuracy	MSE	medianSE	stdSE	SpearmanR
Previous work [17]	Validation	0.379	131549	72900	169116	0.235
Proposed	Validation	0.517	127727	40645	191729	0.429
	Test	0.523	428641	59539	1162454	0.36

greater or equal to the half-width of the border. We finally scale the resulting map so its values are comprised between 0 (far from the boundary) and 100 (on the boundary). The resulting uncertainty map for a given class exhibits a triangular activation shape on the direction perpendicular to the boundary of the objects detected by the segmentation stage.

3 Experiments and Results

Task 1: tumor segmentation Table 1 presents the results obtained for the segmentation task by our proposed method. Overall, we perform better in terms of Dice coefficient on the whole tumor (Dice_WT column) than on the enhanced tumor (Dice_ET, corresponding to the active tumor class) and the tumor core (Dice_TC, corresponding to the necrosis class). Visual inspection on the segmentation infered on the training data set indicated that our method sometimes indeed tends to confuse those two classes together. This is confirmed by the fact that the median Dice is significantly higher than the mean Dice on the whole test data set, implying that eventhough it performs well on most cases, our method seems to really fail in the segmentation of necrosis and enhanced tumor classes on a few cases.
Regarding the Haussdorf distance metric, our method however performs better on the enhanced tumor and tumor core. We believe that this is due to the regularization step performed on the segmented infered by the used FCN architecture, which smoothes the external contours of the edema class, thus impacting the shape of the whole tumor but not the enhanced tumor and tumor core.

Table 3. Performance reported by the automated evaluation platform for uncertainty estimation for our method on the validation data set.

Class	WT	TC	ET
DICE AUC (%) ↑	89.7	72.5	61.7
FTP Ratio AUC (%) ↓	48.3	71.6	70.6

Task 2: survival prediction Table 2 presents the various classification performance metrics, namely the class-based accuracy, the mean, median and standard deviation square errors and Spearman R coefficient for survival predictions expressed in days, for the proposed prediction algorithm for the validation data set and the test data set. For comparison purposes, results obtained by our previous work [17] applied on the validation data set are also presented. The validation and test data sets are comprised of $N = 27$ and $N = 107$ patients, respectively. As it can be seen on Table 2, replacing the conventional PCA (as done in [17]) by the kPCA does indeed improve the class-based classification accuracy since it increases from 0.379 to 0.517 on the validation data set. This accuracy also remains stable when going from the validation data set to the test data set (slightly increasing from 0.517 to 0.523). This notably validates the capacity of kPCA to RFs performances in terms of class-wise prediction with respect to classical PCA. Metrics devoted to the evaluation of the survival prediction in days (namely the mean and median square errors) however do not allow to conclude with respect to the soundness of the final linear regression step for the validation data set when comparing [17] with the currently proposed algorithm.

Task 3: uncertainty quantification As summarized by Table 3, our uncertainty estimation method produces encouraging results for the whole tumor (WT) class as it permits to efficiently filter false positives and increase the overall confidence from an original DICE score of 88.0%, as show by the *DICE AUC* metric. This, however, comes with the price of a large filtering of true positives, as shown by the *FTP Ratio AUC* metric. Regarding the tumor core (TC) and enhancing tumor (ET) classes, the results are less promising as the DICE score degrades from original DICE scores of 74.6% (TC) and 70.1% (ET), mainly because of an over-aggressive filtering. Those results let us believe that this uncertainty estimation method is better suited for cases were the underlying segmentation method already performs quite well. Because of its simplicity and its fast computation, it may be a natural baseline for more complex methods to be compared against.

4 Conclusion

In this article, we present the work submitted for the MICCAI Challenge BraTS 2019, for the segmentation, prediction and uncertaincy tasks.

– The segmentation procedure is performed using two VGG-based segmentation networks;
– The prediction procedure relies on feature extraction and random forests;
– The uncertainty procedure uses the results of the segmentation and gives a confidence score to each pixel based on its distance to the background.

The strength of this method is its modularity and its simplicity. It is easy to implement and fast. From the obtained segmentation result, we propose a simple method to predict the patient overall survival, based on Random Forests, based on the same procedure than during the BraTS 2018 Challenge where we reached the 2nd place. This method only needs as input a segmentation, a brain atlas and a brain volume for atlas registration. It means that our method is robust to the different acquisitions, does not need a special modality or setting, yielding to a method robust to inter-base variations.

References

1. Angelini, E.D., Clatz, O., Mandonnet, E., Konukoglu, E., Capelle, L., Duffau, H.: Glioma dynamics and computational models: a review of segmentation, registration, and in silico growth algorithms and their clinical applications. Current Medical Imaging Reviews **3**(4), 262–276 (2007)
2. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection. The Cancer Imaging Archive **286** (2017)
3. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. The Cancer Imaging Archive **286** (2017)
4. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific data **4**, 170117 (2017)
5. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint arXiv:1811.02629 (2018)
6. Bauer, S., Wiest, R., Nolte, L.P., Reyes, M.: A survey of mri-based medical image analysis for brain tumor studies. Physics in Medicine & Biology **58**(13), R97 (2013)
7. Bonnín Rosselló, C.: Brain lesion segmentation using Convolutional Neuronal Networks. B.S. thesis, Universitat Politècnica de Catalunya (2018)
8. Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., et al.: New response evaluation criteria in solid tumours: revised recist guideline (version 1.1). European journal of cancer **45**(2), 228–247 (2009)
9. Gal, Y., Ghahramani, Z.: Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In: Proceedings of the 33rd International Conference on Machine Learning (ICML-16). pp. 1050–1059 (Jun 2015)

10. Holland, E.C.: Progenitor cells and glioma formation. Current opinion in neurology **14**(6), 683–688 (2001)
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
12. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015)
13. Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W., Kleihues, P.: The 2007 who classification of tumours of the central nervous system. Acta neuropathologica **114**(2), 97–109 (2007)
14. Maninis, K.K., Pont-Tuset, J., Arbeláez, P., Gool, L.V.: Deep retinal image understanding. In: Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Part II. Lecture Notes in Computer Science, vol. 9901, pp. 140–148. Springer (2016)
15. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE transactions on medical imaging **34**(10), 1993 (2015)
16. Ohgaki, H., Kleihues, P.: Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. Journal of Neuropathology & Experimental Neurology **64**(6), 479–489 (2005)
17. Puybareau, E., Tochon, G., Chazalon, J., Fabrizio, J.: Segmentation of gliomas and prediction of patient overall survival: A simple and fast procedure. In: International MICCAI Brainlesion Workshop. pp. 199–209. Springer (2018)
18. Puybareau, E., Zhao, Z., Khoudli, Y., Carlinet, E., Xu, Y., Lacotte, J., Géraud, T.: Left atrial segmentation in a few seconds using fully convolutional network and transfer learning. In: Proceedings of the Workshop on Statistical Atlases and Computational Modelling of the Heart (STACOM 2018), in conjunction with MICCAI. Lecture Notes in Computer Science, vol. 11395, pp. 339–347. Springer (2019)
19. Schölkopf, B., Smola, A., Müller, K.R.: Kernel principal component analysis. In: International conference on artificial neural networks. pp. 583–588. Springer (1997)
20. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR **abs/1409.1556** (2014)
21. Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P.: Random forest: a classification and regression tool for compound classification and QSAR modeling. Journal of chemical information and computer sciences **43**(6), 1947–1958 (2003)
22. Wen, P.Y., Macdonald, D.R., Reardon, D.A., Cloughesy, T.F., Sorensen, A.G., Galanis, E., DeGroot, J., Wick, W., Gilbert, M.R., Lassman, A.B., et al.: Updated response assessment criteria for high-grade gliomas: response assessment in neurooncology working group. Journal of clinical oncology **28**(11), 1963–1972 (2010)
23. Xu, Y., Géraud, T., Bloch, I.: From neonatal to adult brain MR image segmentation in a few seconds using 3D-like fully convolutional network and transfer learning. In: Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP). pp. 4417–4421. Beijing, China (September 2017)